Titre: Title:	Simulation of user perceived QOS in an edge network
Auteur: Author:	Weihua Chen
Date:	2006
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Chen, W. (2006). Simulation of user perceived QOS in an edge network [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/7707/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/7707/
Directeurs de recherche: Advisors:	
Programme: Program:	Non spécifié

UNIVERSITÉ DE MONTRÉAL

SIMULATION OF USER PERCEIVED QOS IN AN EDGE NETWORK

Weihua CHEN
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES

Mai, 2006

Your file Votre référence
ISBN: 978-0-494-17935-2
Our file Notre référence
ISBN: 978-0-494-17935-2

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

SIMULATION OF USER
PERCEIVED QOS IN AN EDGE NETWORK

présenté par: **Weihua CHEN**

en vue de l'obtention du diplôme de: Maîtrise ès science appliquées

a été dûment acceptée par le jury d'examen constitué de:

M.Christian CARDINAL, Ph,D., président

Mme.Brunilde SANSÒ, Ph.D., directrice de recherche

M.Chahé NERGUIZIAN, Ph.D., membre

Acknowledgements

I would like to express my gratitude to all those who gave me the possibility to complete this master thesis.
I am greatly indebted to my thesis supervisor, Professor Brunilde Sansò, for her help, guidance and encouragement during the course of my master research. Her comments have been of greatest help at all times. I am thankful to her for having given me an opportunity to work with her and her other students. I thank her as well for providing me the financial support during my study.
I would like to thank to Professors Christian Cardinal and Chahé Nerguizian for giving me very useful advice as members of the jury. I express sincere gratitude to my colleague in LORLAB, Frédéric Bodevin, for providing me the results and code of GPRS simulator.
My thanks also go to my fellow group members in LORLAB for their encouragement and friendship. I particularly thank Yassine Boujelben and Christian Awad for helping me to translate the condensé en Français in this thesis.
Last but not the least, I would like to thank my husband, my babies, my parents and my brother for being the loving source throughout the years.

Résumé

Les systèmes de la Deuxième-generation (2G) ont popularisé des services de téléphonie sans fils. On s'attend à ce que les systèmes de la troisième-generation (3G) augmentent les systèmes 2G avec des capacités de débits élevés et de multimédia. Dans ce contexte, EDGE (Enhanced Data Rate for GSM Evolution), devient bien plus important, dû à sa compatibilité de 2G et à ses possibilités pour réaliser des débits élevés et l'éfficacité spectrale en utilisant une nouvelle modulation et un mécanisme efficace de contrôle de qualité du lien (LQC).

En déployant un réseau EDGE, un problème important est de pouvoir évaluer la Qualité de service (QS) vue par l'utilisateur. Notre objectif est de trouver un modèle de simulation pour évaluer la Qualité de Service (QS) usager dans un réseau EDGE. Pour cela, on propose un simulateur d'un réseau EDGE qui nous permet de mesurer la Qualité de Service (QS) usager dans un réseau EDGE. Avec ce simulateur, nous produisons une série de résultats pour étudier les facteurs qui influencent la Qualité de Service usager. Ces facteurs incluent: L'utilisation de Contrôle de la qualité du lien (LQC) et les différents algorithmes de LQC; les classes de Qualité de Service basées sur la taille de message; contrôle d'admission et le nombre de canaux réservés pour la transmission de données. En plus, nous étudions aussi l'influence de différentes méthodes utilisées pour servir les services différenciés. Ce simulateur peut nous aider à trouver une manière d'améliorer la Qualité de Service usager d'un réseau EDGE.

Abstract

Second-generation (2G) systems popularized wireless telephony services. Third-generation (3G) systems are expected to enhance 2G systems with high data rate and multimedia capabilities. In this context, EDGE (Enhanced Data Rate for GSM Evolution), becomes even more important, due to its 2G compatibility and its capability to achieve high data rates (384 kbps) and spectral efficiency by using a new modulation and an efficient Link Quality Control (LQC) mechanism.

When deploying an EDGE network, it is important to be able to assess the Quality of Service (QoS) from the user's point of view. The objective of our work is to find a simulation model to evaluate the user perceived QoS in an EDGE network. In order to fulfil that goal, an EDGE simulator is proposed which permits us to measure the QoS offered to users in an EDGE network. With this simulator, we produces series of results to investigate the factors that influence the QoS seen by user. These factors includes: The use of Link Quality Control (LQC) and different LQC algorithms; the QoS classes based on message size; admission control; the number of channels reserved for the data transmission. In addition, we also study the impacts of different method used to serve differentiated users. The simulator can help us to find a way to improve the user perceived QoS in an EDGE network.

Condensé en Français

Introduction

La communication sans fil se développe très rapidement. Bien que la voix soit toujours le service principal des systèmes mobiles, les besoins de communication des données sur l'interface air augmentent. C'est pour cela que le GSM est rapidement amélioré pour incorporer les nouveaux services multimédia.

EDGE (Enhanced Data Rate for GSM Evolution) a été développé pour permettre la transmission de grandes quantités de données à haut débit. Par une nouvelle technique de modulation et un nouveau schéma de codage, un réseau EDGE permet d'augmenter la capacité et l'efficacité sur l'interface air. Le déploiement d'un réseau EDGE introduit un problème important qui consiste à pouvoir évaluer la Qualité de Service (QS) vue par l'utilisateur. Notre objectif est de trouver un modèle de simulation pour évaluer la QS usager dans un réseau EDGE. Pour l'instant, notre travail se concentre sur la transmission de paquets sur un réseau EDGE (c-à-d, EGPRS).

Fonctionnement de EDGE

EDGE est proposé pour augmenter le débit de transfert des données sans faire des changements majeurs aux réseaux GSM/GPRS existants. Comme pour un réseau GSM/GPRS, la même structure de la trame TDMA (Time Division Multiple Access), le canal logique et la bande de porteuse de 200kHz sont employés. Les modifications principales se produisent seulement dans l'interface radio.

Architecture

Le réseau EDGE est semblable au réseau GSM/GPRS. La Figure 1 répresente l'architecture du réseau concernant la transmission de données dans un réseau EDGE.

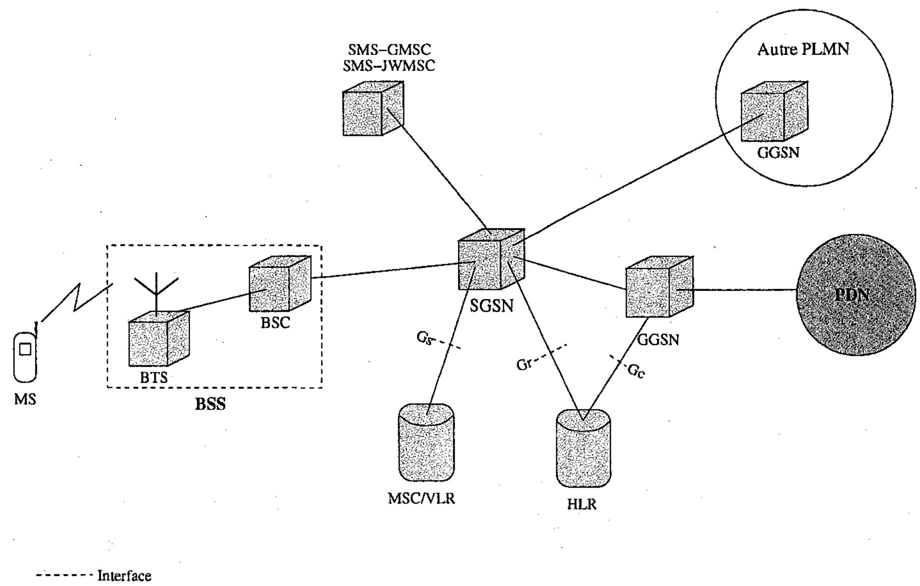


Figure 1: Architecture logique d'un réseau EDGE

Quand l'utilisateur informe le réseau de sa présence et de son désir de commencer une transmission de données, le BSS (sous système de la station de base) coordonne cette demande et informe la MS (station mobile) quelles ressources elle peut employer pour envoyer un message. La MS envoie un message au SGSN (Nœud de services GPRS), ce qui incite le SGSN à informer le HLR (registre des abonnés nominaux) que l'utilisateur est situé dans cette région. Le HLR fournit l'information de profil de service au SGSN. Le SGSN envoie alors un message pour accepter la demande de la MS. La MS doit encore une fois demander les ressources radio du BSS. Quand cette demande est accordée, la MS envoie un message de demande au SGSN. Le SGSN détermine si le service demandé est permis basé sur l'information reçue du HLR. Il détermine aussi quel GGSN (Nœud de support de passerelle) doit être contacté pour fournir le service. Le SGSN fait suivre la demande au GGSN approprié. Le GGSN négocie avec les réseaux externes pour installer le service demandé et informe le SGSN sur quelle information peut être nécessaire pour accomplir la transaction de service. Le SGSN garde l'information appropriée et informe le BSS de tous les détails concernant le trafic suivant. Enfin le SGSN envoie un message d'acceptation à la MS et la session de données peut commencer.

Protocole de transmission de données

Le protocole de transmission de données dans un réseau EDGE est illustré à la Figure 2. Ce protocole est identique à celui de GPRS. La région ombragée montre les protocoles qui sont influencés par l'introduction de EDGE. Les protocoles les plus proches de la couche physique sont les plus affectés par EDGE. Il y a aussi quelques modifications mineures au protocole GPRS du système de station de base. Les autres protocoles restent intacts après l'introduction de EDGE.

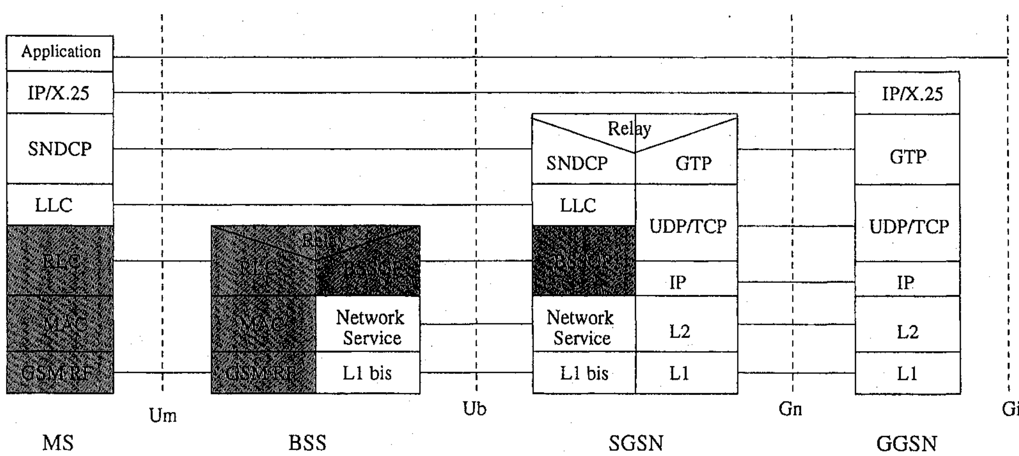


Figure 2: Shéma de transmission d'un réseau EDGE

De la MS au BSS, les données traversent des protocoles fournissant les diverses fonctions pour permettre la transmission de données:

- Le protocole de convergence des sous-réseaux (SNDCP) minimise le transfert de l'information redondante du contrôle (e.g., entête TCP/IP) et des données de l'utilisateur entre le SGSN et la MS par des techniques de compression.

- Les fonctions typiques de LLC comportent chiffrement, le contrôle de flot et le contrôle d'ordre. En outre, si le protocole de LLC est employé dans le mode d'acquittement, il fournit la détection et le rétablissement des erreurs de transmission; en mode de non-acquittement il signale des erreurs irrémédiables.

- La couche de protocole RLC/MAC est responsable de la transmission de lien radio et de l'allocation des ressources radio. Il y a deux concepts importants qui constituent la base de fonctionnement de RLC/MAC: TBF (temporary block flow) et TFI (temporary flow identifier).

Dans le réseau cœur, entre le BSS et le SGSN, le protocole de sous-système de station de base GPRS (BSSGP) transfère des données et fournit de l'information de contrôle. Entre le SGSN et le GGSN, le Protocole d'encapsulation GPRS (GTP) est employé pour transférer les données de l'utilisateur et la signalisation entre les nœuds de support du GPRS (GSN). Tous les paquets sont encapsulés par le protocole GTP.

Modulation

La technique de modulation d'un réseau GSM/GPRS est de type GMSK (Gaussian minimum shift keying). A l'inverse, Dans un réseau EDGE, en plus de la modulation de GMSK, 8PSK (8-phase shift keying) peut aussi être empolyée afin d'atteindre des débits plus élevés.

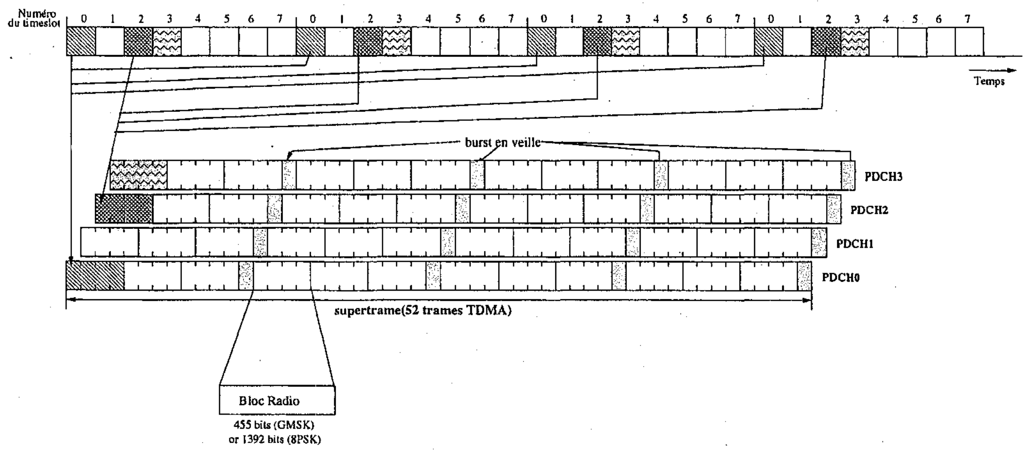


Figure 3: PDCH and 52 TDMA frames structure

Canaux logiques

EDGE utilise la même structure TDMA que le GSM pour former les canaux physiques. Chaque sous bande de 200 KHz est découpée en 8 time-slots de 0.7769 ms qui contiennent 156.25 périodes de symbole, ce qui donne des trames TDMA de 4.613 ms. Chaque time-slot peut être assigné à la transmission de données en mode commutation de paquets, ou en mode commutation de circuits. La récurrence d'un time-slot définit un canal physique. Le canal physique dédié au trafic de paquets de données est appellé un canal de paquets de données (PDCH). La plus petite unité de transmission d'un PDCH est appellé un bloc radio. Quatre time-slots dans quatre trames TDMA consécutives sont utilisées pour transmettre un bloc radio. Comme on peut voir dans la Figure 3, tous les bursts de TS0 appartiennent au PDCH 0. Un PDCH est structuré dans les multi-trames comportant 52 trames TDMA (supertrame), qui correspond à une durée de 240 ms. Le burst en veille n'est pas employé pour transmettre des données, laissant 12 blocs radio dans un supertrame. Ainsi, le temps moyen de transmission par bloc radio est 20 ms.

Selon le type de message transmis dans un bloc radio, une série de blocs radio forme un canal logique (c-à-d, chaque PDCH peut porter plusieurs canaux logiques). Un exemple est un canal de trafic de paquets de données (PDTCH) transportant les données de l'utilisateur.

Schéma de codage

Pour un réseau EGPRS, neuf schémas de codage sont possibles, de MSC1 à MSC9. Les quatre premiers schémas de codage utilisent une modulation de type GMSK alors que les cinq suivants utilisent quant à eux, une modulation de type 8PSK (voir la table 1).

Table 1: Schéma de codage dans un réseau EGPRS

Schéma de codage	modulation	L'infomation dans un bloc radio	Taux de codage			famille	Débit kbits/s
			R_1	R_{1+2}	R_{1+2+3}		
MCS-1	GMSK	176	0.53	0.26	-	C	8.8
MCS-2		224	0.66	0.33	-	B	11.2
MCS-3		296	0.85	0.42	0.28	A	14.8
MCS-4		352	1	0.5	0.33	C	17.6
MCS-5	8-PSK	448	0.37	0.19	-	B	22.4
MCS-6		592	0.49	0.24	-	A	29.6
MCS-7		2×448	0.76	0.38	0.25	B	44.8
MCS-8		2×544	0.92	0.46	0.31	A	54.4
MCS-9		2×592	1	0.5	0.33	A	59.2

Ces neuf codages sont divisés en trois familles: la famille A, B et C. Chaque famille a une unité de charge utile différente contenue dans un bloc radio: 37 (et 34),28 et 22 octets respectivement. La division permet de retransmettre un bloc radio en erreur en utilisant un codage plus robuste de la même famille. Ceci donne la possibilité de la re-segmentation dans un réseau EDGE.

Contrôle de la qualité du lien

Le contrôle de la qualité du lien (LQC) est un terme employé pour indiquer les techniques qui permettent un choix d'un codage selon la qualité courante du lien. Le but du LQC est de maximiser le débit en adaptant la robustesse du lien à la qualité variable du canal.
Dans un réseau EDGE, LQC inclut l'adaptation de lien (LA) et la redondance incrémentale (IR).

- **L'adaptation de lien (LA):** l'adaptation de lien consiste à sélectionner le schéma de modulation et de codage le mieux adapté aux conditions radio rencontrées par le mobile. En cas de conditions favorables, un ensemble modulation-codage efficace est utilisé, même s'il offre une protection réduite. Dans des conditions difficiles, un schéma de modulation-codage robuste est préférable.

- **Redondance incrémentale (IR):** en mode IR, la première tentative de transmission d'un bloc RLC est effectuée avec un codage de faible protection (e.g., P1). Si le décodage est réussi, un débit très élevé peut être obtenu. En cas de mauvaise réception, seulement l'information redondante sera envoyée à la prochaine retransmission en utilisant un schéma de poinçonnage différent. Les blocs en erreurs sont gardés en mémoire (tandis qu'en LA, le bloc en erreur est jeté) et sont combinés avec les informations reçues. Le décodage conjoint avec le bloc initial améliore fortement la chance de décodage réussi, ce qui diminue le taux d'erreur de bloc (BLER).

Modèle de la simulation

On peut voir le modèle de la simulation d'un réseau EDGE dans la Figure 4.

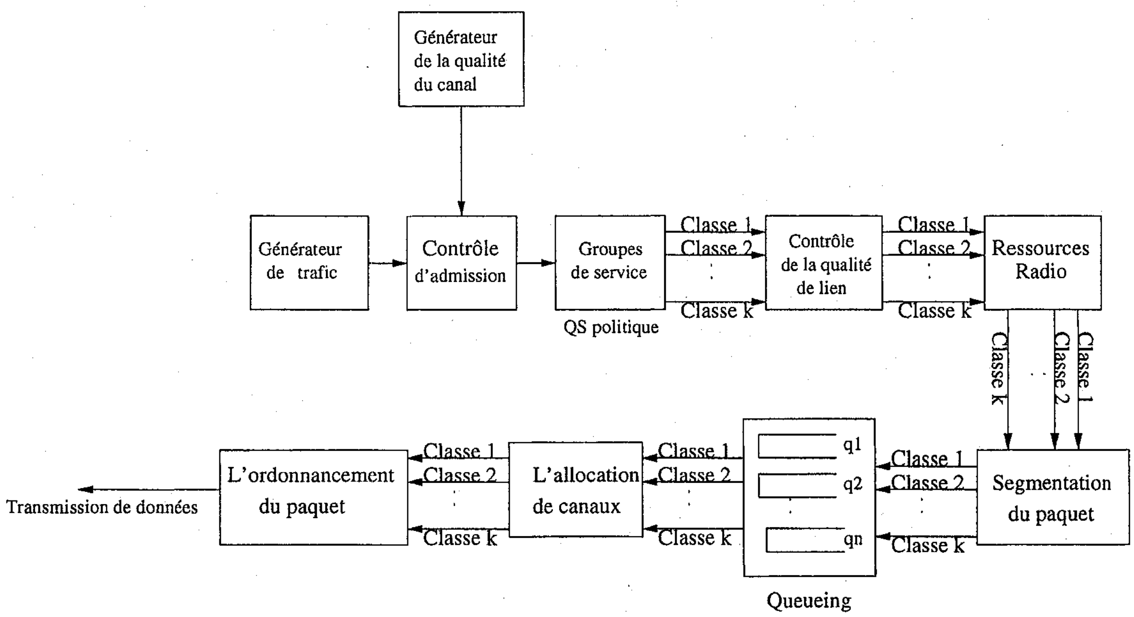


Figure 4: Modèle de la simulation d'un réseau EDGE

Générateur du trafic et de la qualité du canal

Le générateur du trafic produit le trafic de voix et de données, tandis que le générateur de la qualité du canal produit un paramètre pour décrire la condition courante du canal.

Contrôle d'admission

Le but du contrôle d'admission est d'accepter ou non les nouveaux utilisateurs dans le réseau d'accès radio pour améliorer la performance du système. Dans notre simulateur, le contrôle d'admission est basé sur la mesure de la qualité du lien radio. En fait, la transmission avec de mauvais états du lien conduirait à long délai, dégraderait en outre la performance du système. Par conséquent, le refus de la transmission avec de mauvais états du lien améliorerait relativement la qualité de service usager.

Groupes de service

Les paquets d'arrivés peuvent être classifiés en différents groupes de service selon la politique courante de la qualité de service (QS), et les politiques suivantes de QS sont proposées:

- **le meilleur effort:** tous les usager ont la même priorités.
- **Qualité de service basée sur la taille de message:** les classes de QS sont basées sur les tailles de message et peuvent être modélisées par WFQ (Weighted Fair Queuing), qui donnera plus de priorité aux messages courts.
- **Qualité de service basée sur les utilisateurs différenciés:** la différentiation est basée sur les classes du trafic. La qualité de service est fournie aux utilisateurs différenciés, sans tenir compte de leur taille de message.

Ressources Radio

Pour l'instant, nous utilisons un modèle avec une seule bande de fréquence par cellule. Typiquement, sur cette bande, un canal physique est réservé pour la signalisation, trois canaux sont réservés pour la voix, deux canaux sont réservés pour les données et les deux restant peuvent être utilisés pour la voix ou les données avec priorité préemptive pour la voix (voir Figure 5).

Contrôle de la qualité du lien

Les algorithmes suivants de LQC seront employés dans notre simulateur:

- **Un schéma de l'adaptation de lien**

  L'adaptation de lien est mise en application en choisissant un schéma de codage

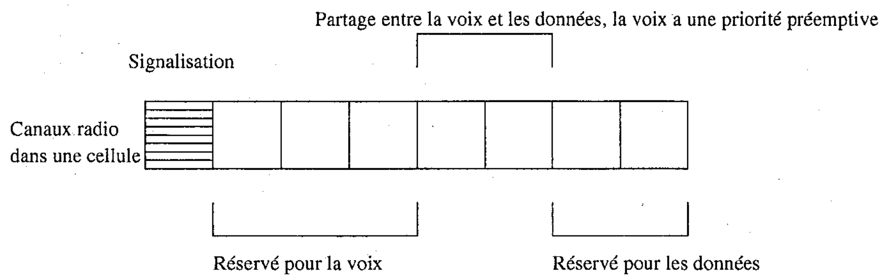


Figure 5: Partage des ressources radio dans une bande entre la voix et les données

réduisant au minimum le délai moyen du transfert T_{delay} décrit en Eq. (1) en veillant à ne pas augmenter excessivement le nombre de retransmissions:

$$T_{delay} = \frac{L}{R_{MCSx} * (1 - BLER_{MCSx})} \tag{1}$$

où L est la taille de paquet IP, R_{MCSx} le débit binaire d'un codage donné, et $BLER_{MCSx}$ le taux d'erreur de bloc de ce codage à un C/I (le rapport entre le signal utile et l'interférence) donné sans combinaison de codage.

- **Un schéma de combinaison de LA et IR**

  En mode de combinaison de LA et IR, le bloc radio est d'abord transmis en utilisant le codage choisi par LA et ensuite, le mécanisme IR a lieu. Lorsque la transmission du bloc RLC n'est pas réussie, seulement l'information redondante sera envoyée par la prochaine retransmission en utilisant un schéma de poinçonnage différent. Dans ce cas-ci, la décision du codage est de minimiser le délai moyen de transfert T_{delay}, décrit en Eq.2, sans augmenter excessivement le nombre de retransmissions.

$$T_{delay} = \frac{L}{R_{MCSx} * (1 - BLER_{eff})} \tag{2}$$

avec $BLER\_eff$ le taux d'erreur efficace de bloc:

$$BLER\_eff \approx \frac{BLER_{MCSx}}{1 + BLER_{MCSx}} \tag{3}$$

Mise en file d'attente

Dans un réseau EDGE, chaque nouvelle transmission demande un TBF (Flot Temporaire de Blocs) d'être installé dans le RLC/MAC, et alors le TBF est attribué les ressources radio sur un ou plusieurs canaux physiques. Il est possible de voir un TBF comme une file d'attente de TBF, qui contient seulement les blocs appartenant à un même message. Une file d'attente de TBF est temporairement maintenue pour la durée la transmission de données. Nous avons choisi de limiter à 4 le nombre de messages qui peuvent être simultanément transmis dans le système, c-à-d, il y a un maximum de 4 files d'attente de TBF, chacune contiente les blocs appartenant à un message qui est en train d'être transmis. Nous considérons aussi un petit tampon qui garde les messages d'arrivés s'il n'y a pas de ressources disponible. Ce tampon ne peut contenir qu'un seul message.

L'allocation des ressources

L'allocation des ressources est prise en compte au niveau du bloc, c'est à dire, le protocole RLC. Les messages doivent être découpés en blocs radio. La taille de charge utile d'un bloc radio dépend du codage courant. L'attribution est effectuée au niveau de la supertrame. Une supertrame est composée de 12 blocs radio B0,B1,B2...,B11. Sur chaque supertrame nous définissons un ou plusieurs canaux logiques auxquels sont attribués les blocs radio. À chaque nouvelle supertrame, chaque message se voit attribuer un ou plusieurs canaux logiques (voir Figure 6). Cette technique permet de transmettre plusieurs messages simultanément sur un même canal physique, ou bien de transmettre un même message sur plusieurs canaux afin d'augmenter le débit.

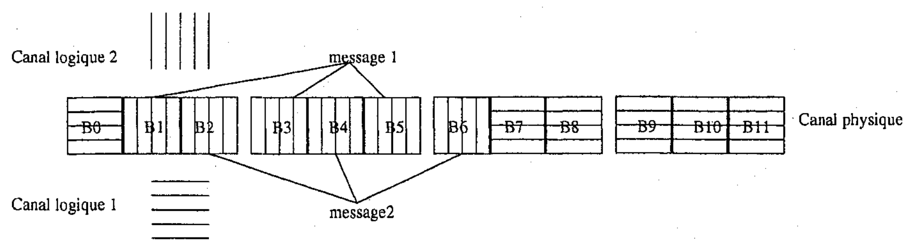


Figure 6: Cannaux logiques et allocation des ressources

L'ordonnancement du paquet

L'algorithme *Weighted Iterative Round Robin (Weighted IRR)* est prise en charge. Toutes les classes qui ont suffisamment de demandes obtiendront les blocs radio d'après les poids associé des classes.

Fournir des services différenciés

Quatre classes de service sont prises en compte dans notre simulateur. Les utilisateurs de la classe 1 ont un poid de 4, les utilisateurs de la classe 2 un poid de 3, les utilisateurs de la classe 3 un poid de 2 et les utilisateurs de la classe 4 un poid de 1. Nous proposons les méthodes suivantes pour fournir le service aux utilisateurs différenciés:

- **File d'attente différencié:** les messages appartenant à la même classe de service ne peuvent pas être transmis en même temps. Les nouvelles arrivées ne seront pas transmises si un message de la même classe est en train d'être transmis. Dans ce système, chaque file d'attente contient seulement les utilisateurs appartennant à la même classe de service. Cependant, chaque file d'attente ne contient qu'un seule message.

- **FCFS (premier-arrivé-premier-servi) sans priorité:** les nouvelles arrivées sont servies avec la politique de premier-arrivé-premier-servi, sans tenir compte de la classe de service. La priorité de la file d'attente de TBF dépend du message contenu dans cette file d'attente.

- **FCFS (premier-arrivé-premier-servi) avec priorité:** cette méthode est semblable à FCFS sans priorité. La nouvelle arrivée est servie selon son temps d'arrivée sans considérer la classe à la laquelle il appartient. Mais les utilisateurs de la classe 1 peuvent occuper les ressources qui sont en train d'être utilisées par les utilisateurs de la classe 4, s'il n'y a plus de ressource disponible. Les utilisateurs de la classe 4 sont jetées même si elles n'ont pas fini ses transmission.

Mesures de la qualité de service usager

La qualité de service est évaluée au niveau du message et nous avons choisi d'étudier les mesures suivantes de la qualité de service: *le temps total de transfert*, *le temps total de transfert du premier bloc de données* et *le blocage.*

Implémentation de la simulation

Les programmes du modèle de simulation sont codés en employant le language C++. La simulation est en partie basée sur la librairie d'une simulation d'évènements discrets écrite par D.Boliers et A.Eliens de la faculté des sciences d'Amsterdam. Cette librairie est disponible à l'adresse http://www.cs.vu.nl/ eliens/sim/.

Résultats de la simulation

Avec notre simulation, plusieurs séries de résultats ont été produites pour trouver les facteurs qui influencent la qualité de service vue par l'usager. Selon ces résultats, l'importance du contrôle de qualité du lien a été vérifiée. Le nombre de canaux réservés aux données est une variable très importante de performance. Nous pouvons évaluer l'amélioration de la qualité de service usager qui résulte d'une augmentation du nombre de canaux réservés aux données et quantifier en même temps la dégradation de performance de la voix. La combinaison de LA et IR peut produire une meilleure qualité de service usager sans influencer la qualité de service de la voix. le contrôle d'admission est une autre technologie qui peut diminuer le temps de transfert de données, mais il pourrait causer un excès de blocage des messages lorsque la charge du système est faible. L'introduction de classes de services permet une meilleure exploitation du réseau et prend mieux en compte les besoins des clients. Nous étudions deux politiques différentes de la gestion de classes de service (classes de service basées sur la taille de message et classes de services basées sur les usagers diffrenciés). Avec la politique basée sur la taille de message, la qualité de service usager peut être améliorée, sans endommager la qualité de service de la voix. Quand aux classes de services par usager, contrairement aux classes de services par longueur de messages où les messages doivent être assortis selon leur longueur, un usager utilise la même classe de service pour tous ses messages, sans tenir compte de leur longueur. Ceci permet de classifier facilement les groupes d'utilisateur. Cette politique améliore la qualité de service des utilisateurs avec priorités élevés en dégradant de manière significative les performances de ceux qui ont ont la basse priorité. Selon les résultats, nous pouvons dire que la file d'attente différenciée et le FCFS avec priorité donnent l'avantage aux deux premiéres classes de services. Par conséquent, nous pouvons employer l'une ou l'autre de la méthode selon la qualité de service que nous voulons garantir.

Enfin, il sera intéressant de modéliser des cellules multiples et tenir compte de trafic de handoff. Un autre travail intéressant est le service IP de temps réel basé comme voix ou vidéo sur IP.

Contents

List of Tables

List of Figures

Results for effects of Link Quality Control

Results for QoS policy based on message size

Results for QoS policy based on Differentiated Services

List of Abbreviations

8PSK	8 Phase-Shift Keying
ARQ	Automatic Repeat Request
BEP	Block Error Probability
BLER	Block Error Rate
BSC	Base Station Controller
BSS	Base-station Sub System
BSSGP	BSS GPRS Protocol
BTS	Base Transceiver Station
CCU	Channel Codec Unit
C/I	Carrier-to-Interference Ratio
ECSD	Enhanced Circuit-Switched DATA
EDGE	Enhanced Data rate for GSM Evolution
EGPRS	Enhanced GPRS
FCFS	First-Come-First-Serve
GERAN	GSM/EDGE Radio Access Network
GGSN	Gateway GPRS Support Node
GMSC	Gateway Mobile Switch Center
GPRS	Generic Packet Radio Service
GMSK	Gaussian Minimum-Shift Keying
GSM	Global System for Mobile communication
GTP	General Telemetry Processor

HLR	Home Location Register
HSCSD	High-Speed Circuit Switched Data
IP	Internet Protocol
IR	Incremental Redundancy
IRR	Iterative Round Robin
IWMSC	Interworking MSC
LA	Link Adaptation
LLC	Logical Link Control
LQC	Link Quality Control
MAC	Medium Access Control
MCS	Modulation and Coding Scheme
MSC	Mobile Switch Center
PACCH	Packet Associated Control Channel
PAGCH	Packet Access Grant Channel
PCU	Packet Control Unit
PDCH	Packet Data Channel
PDP	Packet Data Protocol
PDTCH	Packet Data Traffic Channel
PDU	Protocol Data Unit
PPCH	Packet Paging Channel
PRACH	Packet Random Access Channel
QoS	Quality of Service
RB	Radio Block
RF	Radio Frequency
RLC	Radio Link Control
RR	Radio Resource
RRC	Radio Resource Control
SDNCP	Sub-network Dependent Convergence Protocol

SGSN	Serving GPRS support Node
SMS	Short Message Service
TBF	Temporary Bloc FLow
TCP	Transmission Control Protocol
TFI	Temporary Flow Identity
UMTS	Universal Mobile Telecommunications System
VLR	Visitor Location Register
WRR	Weighted Round Robin

Chapter 1

Introduction

1.1 Development of cellular network system

Wireless communication is developing more rapidly than any other major segment of the telecommunications sector. Although voice is still the main service of the mobile systems, the popularity of data communications over the air interface are increasing. Even if data transfer by circuit mode is a part of the GSM protocol, the data service is difficult to exploit well due to its slowness (9.6 kbps), its long connection time, and its billing based on call duration. So the burst traffic of Internet leads to an ineffective radio resource utilization. That is why GSM has been upgraded to incorporate new multimedia services: GSM phase 2+, HSCSD (High-Speed Circuit Switched Data) and GPRS (General Packet Radio Service) have already been standardized for real-time data and end-to-end packet data services respectively.

The purpose of HSCSD is to provide real time high speed data services using circuit-switching. Unlike GSM in which only one time slot is allocated to a voice call, HSCSD can allocate multiple time slots per data call throughout the duration of communication in order to increase the transmission rate of the air interface. HSCSD is suitable for non-bursty high speed data applications such as video-conference and file transfer.

In contrast with the circuit switched transmission of GSM, GPRS is designed for end-to-end packet oriented services having bursty characteristics such as web browsing.

Both HSCSD and GPRS utilize multi-slot operation to achieve high bit rates. But since both of them use Gaussian Minimum Shift Keying (GMSK) modulation, the increase of bit rates is limited. Consequently, to comply with current data rate requirement [29], a large number of time slots would need to be allocated to each data connection. It is unacceptable because it would significantly impact the communication of voice call due to the scarcity of the available frequency spectrum. Therefore EDGE (Enhanced Data Rates for GSM Evolution) has been proposed to overcome the shortcomings of GPRS and HSCSD.

EDGE introduces a new modulation technique named 8 Phase-Shift Keying (8PSK)[18] and new channel coding[17] that can be used to transmit both packet-switched and circuit-switched voice and data services. Therefore higher bit rates (up to 384 kbps) can be achieved. With this data rates it will be possible to offer 3G services (e.g. movie, music) based on 2G systems. Thus, EDGE is a cost effective solution for 3G service and is the ideal choice for operators that do not own UMTS (Universal Mobile Telecommunication Services) licenses.

The first step of the EDGE standard includes the Enhanced Circuit Switched Data (ECSD) based on HSCSD and the Enhanced General Packet Radio Service (EGPRS) based on GPRS. Thanks to the introduction of 8PSK modulation, link quality control and multi-slot techniques, EGPRS and ECSD can offer significantly higher throughput and capacity.

The second phase of EDGE is known as GERAN (GSM/EDGE Radio Access Network), which is responsible of providing UMTS standards based on the evolved GSM core network. The main objective of GERAN standardization is the support for generic packet based realtime services.

1.2 Thesis objective

EDGE is firstly proposed as a standard. All the technique aspects have already been defined in the standards, e.g. the protocols, the radio link, the Link Quality Control (LQC) algorithms, the profile of quality of service (QoS), etc. However, it may exist differences between the theoretical promises of a protocol and the practical performances. For example, it is difficult to evaluate the performances of the radio link in a general way, the environmental factors being too significant. On the one hand, the equipment suppliers, who are responsible for producing the necessary elements to the construction of a network (bases, mobiles, switches, etc), must adapt the technological standards. In fact, many points of establishment are not defined in the standards but left manufacturer dependent. On the other hand, the operators, who market this network, must rely on the theoretical characteristics of their network to propose realizable and interesting solutions for the consumers. It is clear that it is essential to be able to evaluate the real performances of a network under penalty of proposing unrealistic solutions.

The objective of our work is to propose a suitable simulation model to evaluate the user perceived QoS in an EDGE network. For the time being, our work focuses on the packet data transmission over an EDGE network. The parameters that we have chosen to measure the user perceived QoS are the mean delivery time, the maximum delivery time for 95% of the users, the system response time and the block rate.

1.3 Contribution

The contributions of our work are listed as follows:

- We propose a simulation model that was based on the previous GPRS simulator developed in [32]. The following additional modifications were carried out to assess the QoS offered to the users in an EDGE network:
  - A channel condition generator is added to produce the parameter of current channel condition;
  - Admission control is taken into account to improve the QoS;
  - Link Quality Control is modelled in the EDGE simulator;
  - A new queueing system is proposed to provide differentiated services.
- With this simulator, we analyze the influence of link quality control(LQC) and the admission control.
- We also propose different queueing methods to provide the service to different classes of users and our simulation results can help us find a best method to satisfy the user perceived QoS.

1.4 Organization

This thesis is organized as follows. First of all, we study the functions of an EDGE network, especially its packet data component (i.e. EGPRS) functionality. This is presented in Chapter 2. Then, a literature review will be presented in Chapter 3. The simulation model and its implementation will be proposed in Chapter 4. Simulation results are presented in Chapter 5. Finally, the Conclusions can be found in Chapter 6.

Chapter 2

EDGE Functionality

The purpose of EDGE is to increase user transfer data rate without making major changes to existing GSM/GPRS networks. Just as for GSM/GPRS, the same TDMA (Time Division Multiple Access) frame structure, logical channel and 200kHz carrier bandwidth are used. Therefore, the existing cell plans can be greatly utilized to provide multimedia services.

The major modifications only occur at the radio transmission interface. Figure 2.1[2][15] shows that hardware and software upgrades are needed to support the new air interface modulation and the increased data rates. Here the air interface is the radio-frequency portion of the circuit between the cellular phone set or wireless modem (usually portable or mobile) and the active base station. The voice call initiated by the Mobile Station (MS) is first sent to the Base Station Subsystem (BSS), and then follows the path from the BSS to the Mobile Switch Center (MSC), and on to the Public Switched Telephone Network (PSTN) through the Gateway MSC (GMSC). The traffic data will be sent to the external IP network through the path from the BSS to the GPRS support nodes (SGSN,GGSN).

The higher data rates are achieved by using a different modulation scheme and an efficient Link Quality Control (LQC) technique. No new hardware is required in the core network. Therefore, the implementation of EDGE in the existing mobile networks is simple. Only an extra EDGE transceiver unit to each cell is needed. The new EDGE capable transceiver can also handle standard traffic and automatically switch to the EDGE mode when needed.

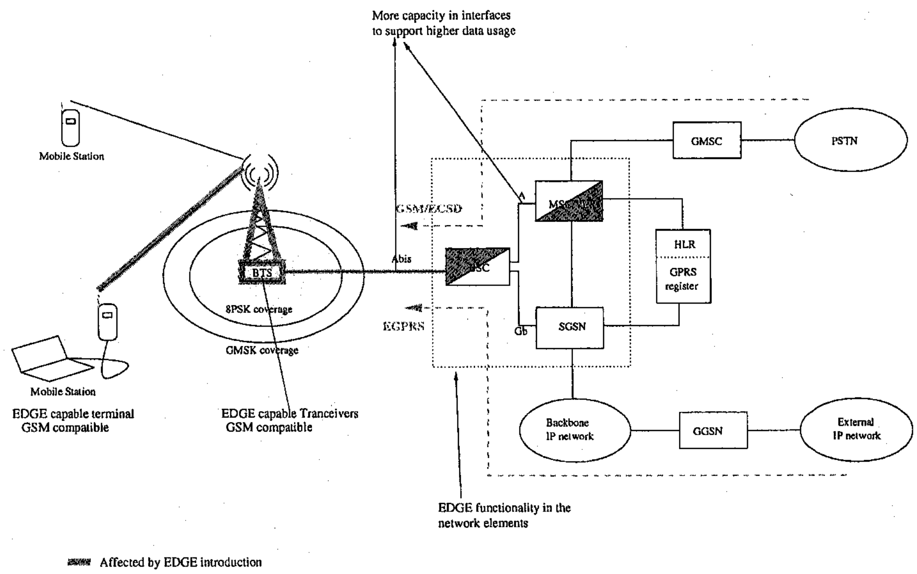


Figure 2.1: EDGE in GSM/GPRS cited from [2]

2.1 Network Architecture

The EDGE network is similar to the GSM/GPRS network. Figure 2.2 depicts the EDGE network architecture (described in [25]) relevant to the packet transmission. The major elements in this architecture are: the Mobile Stations (MS), the Base Station Subsystem (BSS), the Serving GPRS Support Node (SGSN), the Gateway GPRS Support Node (GGSN), the Home Location Register (HLR), the Mobile Switch Center (MSC)/visitor location register (VLR) and SMS-GMSC (Short Message Service Gateway MSC).

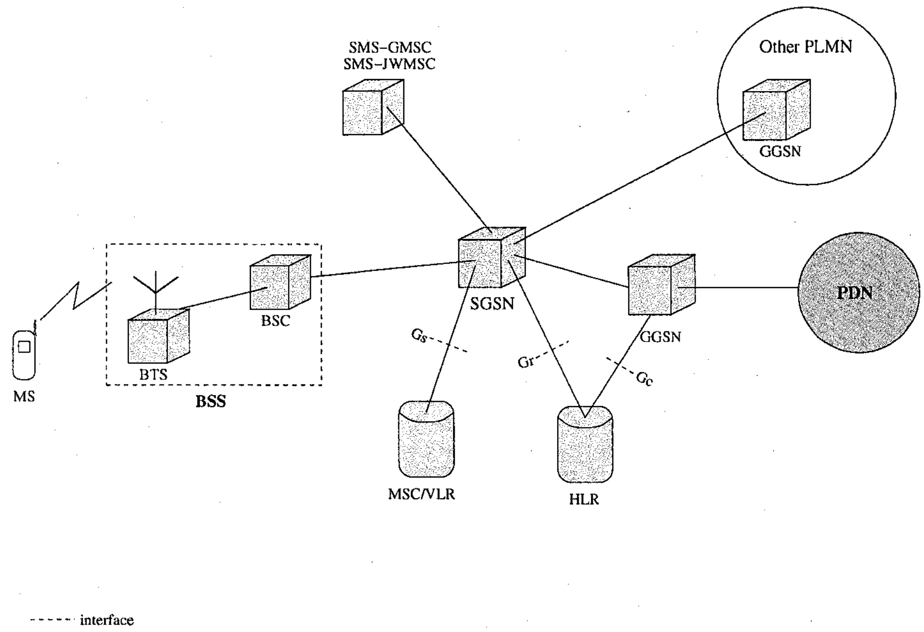


Figure 2.2: EDGE logic architecture [25]

2.1.1 Mobile Stations (MS)

GSM mobile stations must be designed with the appropriate protocol layers for them to support GPRS or EDGE. They must be modified to operate on shared traffic channels and the coding schemes defined in EDGE standard must be added. If the MS is EDGE capable, it must also implement a new modulation scheme (8PSK).

2.1.2 Base Station Subsystem (BSS)

The BSS consists of the Base Transceiver Station (BTS) and the Base Station Controller (BSC)[26]. The BTS is comprised of all the radio transmission and reception equipment. It performs the modulation and demodulation. The BSC, including the Packet Control Unit (PCU), supports all relevant (E)GPRS protocols for communication over air interface. The PCU's function is to set up, supervise and disconnect packet-switched calls, including support for cell change, radio resource configuration and channel assignment.

2.1.3 Serving GPRS Support Node (SGSN)

The SGSN handles mobility functions and controls the data flow toward the BSC. The SGSN provides a point of attachment for the (E)GPRS mobiles. Once the mobile station has attached to the system, a logical link is established between the mobile station and the SGSN, via the base station. The tasks of the SGSN includes session management, mobility management, and logical link management to the MS. It also provides a connection to the database, such as the HLR, in the MSC.

2.1.4 Gateway GPRS Support Node (GGSN)

The GGSN provides connectivity to the external packet data networks (PDN). The primary role of the GGSN is to route data to the mobile stations at their current points of attachment. All packets between the external PDNs and the GPRS network enter and exit from the GGSN. Once the mobile station activates its packet data address, the MS is registered with the corresponding GGSN. The GGSN then assigns the correct SGSN for the MS according to the location of the MS.

2.1.5 GSM functional entities

The MSC/VLR, HLR, and SMS-GMSC/SMS-IWMSC are functional entities of the initial circuit-switched GSM. These nodes are enhanced by additional interfaces for interworking with GPRS. The connection between SMS-GMSC/SMS-IWMSC and the SGSN enables the (E)GPRS MS to send and receive small messages over

(E)GPRS radio channels. The HLR contains subscribed QoS profile of (E)GPRS MS and routing information. The HLR is accessible from the SGSN via the Gr interface and from the GGSN via the Gc interface. The MSC/VLR is connected with SGSN via Gs interface. Therefore, paging for circuit switched calls can be performed more efficiently via the SGSN.

Here is the data transfer process through these network nodes. When the user informs the network of its presence and its desire to initiate a data transmission, the Base Station Subsystem coordinates this request and notifies the mobile station about the resources that can be used to send a message. The mobile sends a GPRS attach message to the SGSN, which triggers the SGSN to perform authorization, to check authentication and to notify the Home Location Register that the user is located in this SGSN service area. The HLR provides service profile information to the SGSN so it can coordinate the service request. The SGSN then sends a message to accept the requirement of the mobile station. The mobile station must activate its packet data profile (PDP) before it can exchange any data. The MS must, once again, request radio resources from the BSS to begin the PDP context activation process. When this request is granted the mobile station sends a request message to the SGSN. The SGSN determines if the request service is allowed based on the service profile information received from the HLR. It also determines which GGSN needs to be contacted to provide the service that was requested by the user. The SGSN forwards the request to the appropriate GGSN. The GGSN negotiates with external networks to set up the requested service and informs the SGSN the PDP address for the mobile and any additional information that may be necessary to complete the service transaction. The SGSN stores the relevant information and notifies the BSS of any specifics regarding subsequent traffic. Finally the SGSN sends an accept message to the mobile station and the data session may begin.

2.2 Protocol architecture

The EDGE data transmission protocol is illustrated in Figure 2.3[25]. This is the same as the GPRS protocol stack. The shaded area shows the protocols that are influenced by the introduction of EDGE. The protocols closest to the physical layer (the radio link control and the media access control) are most affected by EDGE (see [19]). There are also some minor modifications to the base station system GPRS protocol. In addition to these changes, the rest of the protocol stack remains intact after the introduction of EDGE.

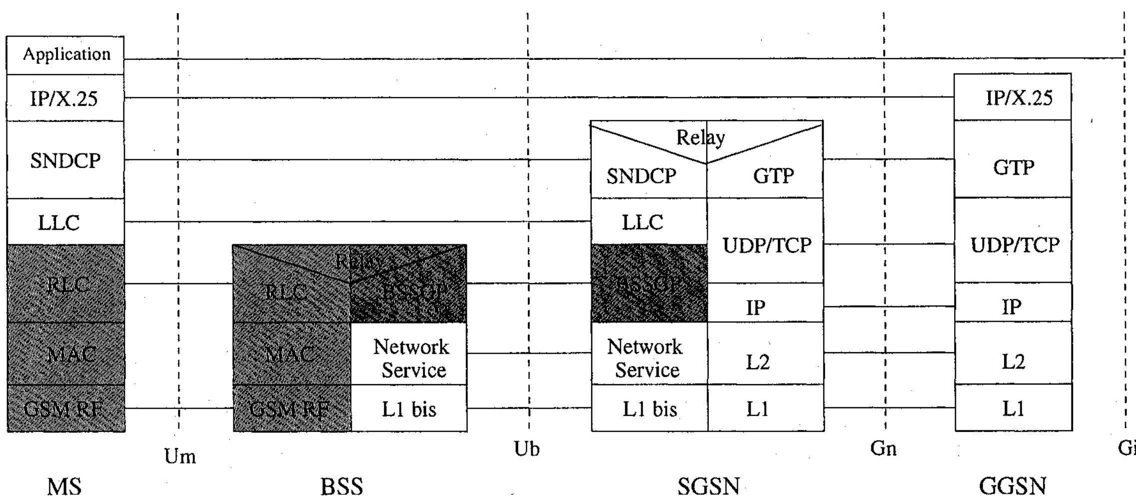


Figure 2.3: Transmission plane protocol architecture

In the BSS, the protocol stack has been broken in order to show how the different entities may be deployed in the radio access network. The channel codec unit (CCU), i.e. the layer 1 functions, is in the BTS; whereas the GSM packet control unit,i.e. medium access control (MAC) and radio link control (RLC) functions, is deployed in the BSC[25][27].

From the MS to the BSS, data is passing through protocols providing various functions for enabling data transmission. First, applications run on top of IP/X.25, which is the packet protocol offered to the subscriber by the (E)GPRS system.

Then descending in the protocol stack, the sub network dependent convergence protocol (SNDCP) minimizes the transfer of redundant control information (e.g.TCP/IP header) and user data between the SGSN and the MS through compression techniques. The output of the compression subfunctions are segmented into LLC (logical link control) frames[25].

The LLC protocol provides a logical link between the MS and its SGSN. Typical LLC functions comprise ciphering, flow control and sequence control. In addition, if the LLC protocol is used in the acknowledged mode, it provides detection and recovery of transmission errors; in the unacknowledged mode it signals unrecoverable errors. The LLC is used by SNDCP for the transfer of network layer packet data units (PDUs)[20].

The Radio Link Control/Medium Access Control (RLC/MAC) protocol layer located in the PCU provides services for the transfer of LLC PDUs using a shared medium between multiple MS and the network[18].

The functions RLC (Radio Link Control) includes the segmentation of LLC PDUs (Packet Data Unit) into RLC data blocks and reassembly of RLC data blocks into LLC PDU; backward error correction (BEC) enabling the selective retransmission of RLC data blocks[19]. The MAC protocol realizes the different logical channels needed to share common transmission medium by several MSs[19]. It allows one MS to use several physical channels (i.e. time slots) in parallel, but also the multiplexing of several MSs over one physical channel.

The physical radio frequency (RF) layer performs transmission and reception of modulated waveforms on the carrier frequencies and is identical to the traditional GSM RF layer.

Within the core network, between the BSS and the SGSN, the Base Station System GPRS Protocol (BSSGP) transfers data and provides control information. Below the BSSGP, the Network Services have load sharing and redundancy functions[24]. Between the SGSN and the GGSN, the GPRS Tunnelling Protocol (GTP) is used to tunnel data and signaling between the GPRS support nodes. This is the process of adding a header to the existing packet so that it can be routed through the backbone network. When the packet reaches the far side of the GPRS network the additional header is discarded and the packet continues on its route based on the original header. The use of tunnelling helps solve the problem of mobility for the packet networks and eliminates the complex task of protocol interworking.

2.3 Air interface

The introduction of EDGE induces the modifications of air interface to facilitate bit rates higher than those that can be obtained from current cellular system. In addition, the air interface (i.e. Um interface)is considered one of the central aspects of EDGE, because it mainly determines the performance of an EDGE network. The air interface is explained in more detail in the following subsections.

2.3.1 Modulation

The modulation used in GSM/GPRS is the Gaussian minimum shift keying (GMSK) [22], which is a kind of phase modulation. In the EDGE system, in addition to the GMSK modulation, 8PSK (8-phase shift keying) scheme can also be employed in order to achieve higher bit rates[22]. Figure 2.4 is the signal constellation of these two modulations, where I means in-phase component of the signal and Q means quadrature.

For GMSK, Every symbol that is transmitted represents one bit; that is, each shift in the phase represents one bit. In GMSK modulation scheme, the signal is passed through a Gaussian filter for pulse-shaping, which gives it a constant modulation envelope[35]. The constant envelope obtained in GMSK makes it more robust in multipath fading environments. In addition, in case there is a phase error, only one

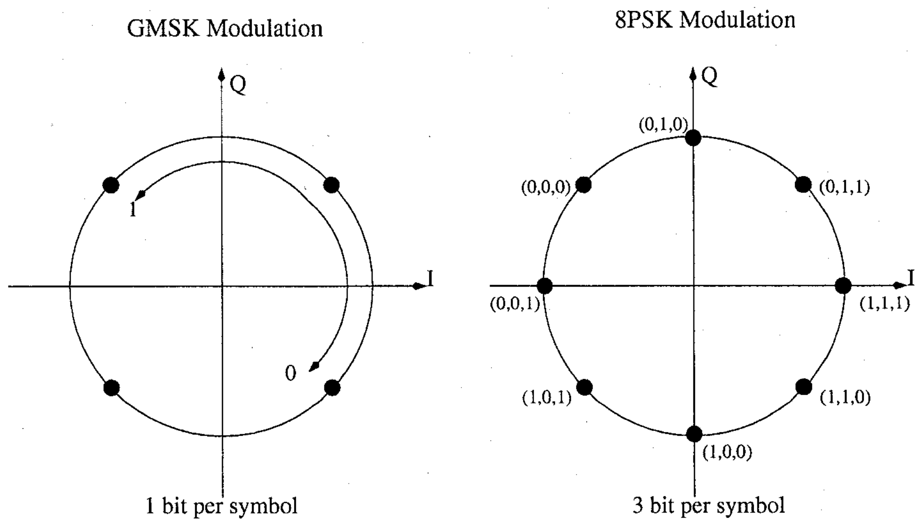


Figure 2.4: Signal constellations for 8PSK and GMSK[22]

bit will be lost, which makes GMSK a robust modulation scheme. On the other hand, the transmission rate of GMSK is relatively low. 8PSK modulation has the same qualities in terms of generating interference on adjacent channels as GMSK. This makes it possible to integrate EDGE channels into an existing frequency plan and to assign new EDGE channels in the same way as standard GSM channels. The 8PSK modulation method is a linear method in which three consecutive bits are mapped onto one symbol. The symbol rate remains the same as for GMSK, but each symbol now represents three bits instead of one. The total data rate is therefore increased by a factor of three [22]. But the transmission reliability of 8PSK is lower than that of GMSK because of its amplitude variation and higher bit loss in case of phase error. Therefore, in an EDGE network, 8PSK and GMSK can coexist. With 8PSK it is possible to provide higher data rates with a reduced coverage, whereas GMSK will be used as a robust mode for wide area coverage.

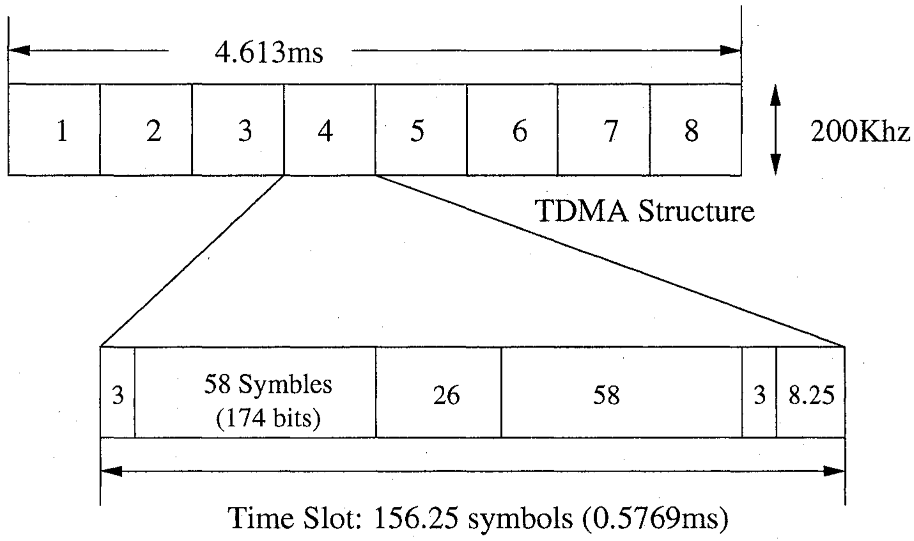


Figure 2.5: EDGE TDMA and Timeslot Structure

2.3.2 Logical channels

EDGE uses the same TDMA structure as GSM to form physical channels[21]. Figure 2.5 portrays the TDMA and time slot structure in an EDGE network[16]. The 200kHz carrier bandwidth is divided into TDMA frames with a length of 4.613 ms. Each TDMA frame is further split up into 8 time slots of equal size. Each time slot can be assigned to either packet-switched data transmission, or circuit-switched data transmission. A time slot is divided into 156.25 symbol periods ($577\mu s$), including a training sequence of 26 symbols in the middle, three tail symbols at either end, and 8.25 guard symbols at one end. Each burst carries 2×58 data symbols, each comprising 3 bits.

The repetition of a definite time slot defines a physical channel. The physical channel dedicated to packet data traffic is called a Packet Data Channel (PDCH). The basic transmission unit of a PDCH is called a radio block. Four time slots in four consecutive TDMA frames compose a radio block. As can be seen in Figure 2.6 [18][21], all bursts of TS0 belong to PDCH 0. A PDCH is structured in multiframes comprising

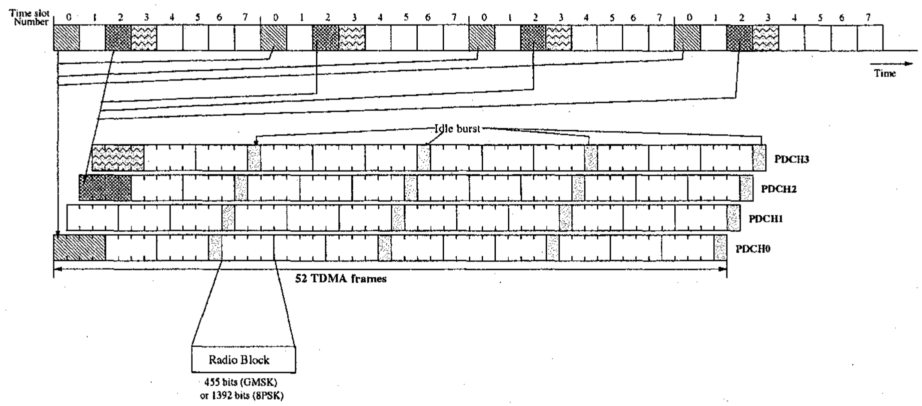


Figure 2.6: PDCH and 52 TDMA frames structure

52 TDMA frames, which corresponds to a duration of 240 ms. The idle burst is not used to transmit data, leaving 12 radio blocks in one multiframe. Thus, the mean transmission time per radio block is 20 ms. Every coded radio block has a length of 456 bits when using GMSK or 1392 bits when using 8PSK, but due to forward error correction fewer payload bits can be transmitted. The number of payload bits contained in a radio block depends on the modulation and coding scheme[18].
Depending on the message type transmitted in one radio block, a sequence of radio blocks forms a logical channel (i.e. each PDCH can carry several logical channels). An example is a Packet Data Traffic Channel (PDTCH) transporting user data.
Some of the logical channels are briefly described as follows[18]:

- **Packet random access channel (PRACH),uplink**: Common channel used by the MSs to initiate an uplink transfer.

- **Packet paging channel (PPCH),downlink**: The BSC uses this channel to page MSs prior to downlink data transmission.

- **Packet access grant channel (PAGCH),downlink**: Resource assignments for up- and downlink transfers are sent on this channel.

- **Packet data traffic channel (PDTCH),up- and downlink:**Data packets are sent on this channel. An MS can use one or several PDTCHs.

- **Packet associated control channel (PACCH),up- and downlink:**This channel conveys signaling information related to a given MS and the corresponding PDTCHs (e.g., RLC acknowledgments).

2.3.3 Medium Access Control and Radio Link Control

The Medium Access Control (MAC) and Radio Link Control (RLC) layer operates above the physical link layer in the protocol architecture. The MAC is responsible for configuring the mapping of logical channels onto the appropriate physical channels. The MAC function defines the procedures for cell selection and re-selection, resource allocation (queuing and scheduling of access attempts) and the provision of Temporary Block Flows (TBFs) that allow point-to-point transfer of data within a cell between the network and a mobile station. A reservation protocol based on the Slotted Aloha protocol is used for contention resolution among several mobile stations. The MAC layer aids in queuing and scheduling of the access attempts.

The RLC function defines the procedures for a selective re-transmission of unsuccessfully delivered RLC data blocks[19]. The basic principle of data transfer is illustrated in Figure 2.7[19]. The IP packets are compressed and segmented into the sub-network protocol data units (SN-PDU) by the SNDCP protocol. IP-packet compression is optional. The SN-PDUs are encapsulated into one or several LLC frames. The size of the data part of the LLC frames is a parameter between 140 and 1520 bytes. LLC frames are segmented into RLC data blocks. At the RLC/MAC layer, a selective automatic repeat request (ARQ) protocol (including block numbering) between the MS and the network provides re-transmission of erroneous RLC data blocks. When a complete LLC frame is successfully transferred across the RLC layer, it is forwarded to the LLC layer.

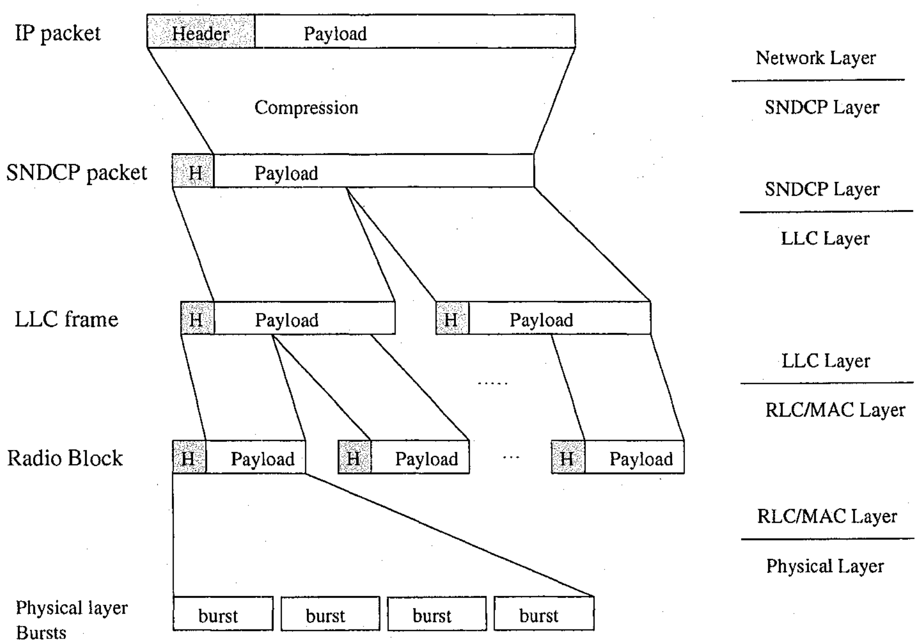


Figure 2.7: Transmission and reception data flow

There are two important concepts that constitute the core of RLC/MAC operation: Temporary Block Flow (TBF) and Temporary Flow Identifier (TFI)[18].

- Temporary Block Flow

  A Temporary Block Flow (TBF) is a physical connection between the MS and the network to support the uni-directional transfer of LLC PDUs on packet data physical channels. A TBF can use radio resources on one or more PDCHs and comprises a number of RLC/MAC blocks carrying one or more LLC PDUs. A TBF is temporary and is maintained only for the duration of the data transfer.

- Temporary Flow Identity

  Each TBF is assigned a Temporary Flow Identity (TFI) by the network. The TFI is assigned in a resource assignment message that precedes the transfer of packets belonging to one TBF to/from the MS. The assigned TFI is unique among concurrent TBFs in each direction. The same TFI is included in every RLC header belonging to a particular TBF, as well as in the control messages associated with the LLC frame transfer (e.g. acknowledgements) in order to address the peer RLC entities.

To support the packet-switched principle of (E)GPRS, the resources of one PDCH are assigned only temporarily to one MS. The BSC controls resources in both the downlink and uplink directions. Once an MS is successful with its channel request and the relative resources are available, a TBF will be established. When a TBF is established, resources (e.g.PDTCH) are assigned to an MS and data transmission can start. As soon as all data for one MS is successfully transmitted, the TBF is released. The duration of TBF ranges from some milliseconds up to several minutes, depending on the amount of data that need to be transmitted.

2.3.4 Channel Coding

By using both 8PSK and GMSK modulation, nine channel coding schemes are defined in EGPRS packet data channel[17]. EDGE modulation and coding schemes are shown in table 2.1[18].
In this table, the raw data with in one radio block (RB) is the transmitted information contained in one radio block, e.g. for MCS-6, one radio block contains 592 bits transmitted information. R_1,R_{1+2} and R_{1+2+3} are the code rate when using an incremental redundancy in EDGE retransmission. R_1 is the initial code rate when an RLC block is first transmitted. If this block cannot be decoded correctly, more redundancy bits will be retransmitted, resulting in a lower code rate R_{1+2}. And so does R_{1+2+3}. Details will be described in section 2.3.5.
These nine MCSs are divided into three families[18]: Family A, B and C (see Figure 2.8). Each family has a different unit of payload contained in a radio block: 37

Table 2.1: EGPRS Modulation and Coding Scheme

small Coding Scheme	Type of modulation	RLC Block per RB	Raw Data within one RB	Code Rate			family	Data rate kbits/s
				R_1	R_{1+2}	R_{1+2+3}		
MCS-1	GMSK	1	176	0.53	0.26	-	C	8.8
MCS-2		1	224	0.66	0.33	-	B	11.2
MCS-3		1	296	0.85	0.42	0.28	A	14.8
MCS-4		1	352	1	0.5	0.33	C	17.6
MCS-5	8-PSK	1	448	0.37	0.19	-	B	22.4
MCS-6		1	592	0.49	0.24	-	A	29.6
MCS-7		2	2×448	0.76	0.38	0.25	B	44.8
MCS-8		2	2×544	0.92	0.46	0.31	A	54.4
MCS-9		2	2×592	1	0.5	0.33	A	59.2

(and 34), 28 and 22 octets respectively. For example, for MCS-9, the transmitted information contained in a radio block is 1184 bits ($= 4 \times 37 octets$); for MCS-6, a radio block contains 592 bits ($= 2 \times 37 octets$) information; for MCS-3, the information contained in a radio block is 296 bits (37octets). MCS-9,MCS-6 and MCS-3 belong to family A.

This division makes it possible to retransmit an erroneous RLC block using a more robust coding scheme from the same family [19]. For example, if MCS-7 is selected for the first transmission of an RLC block, any MCS of family B (MCS-7,MCS-5,MCS-2) can be used for the retransmission if necessary. It is called the capability of resegmentation in EDGE (see Figure 2.9).

Figure 2.10 gives an example of resegmentation process.

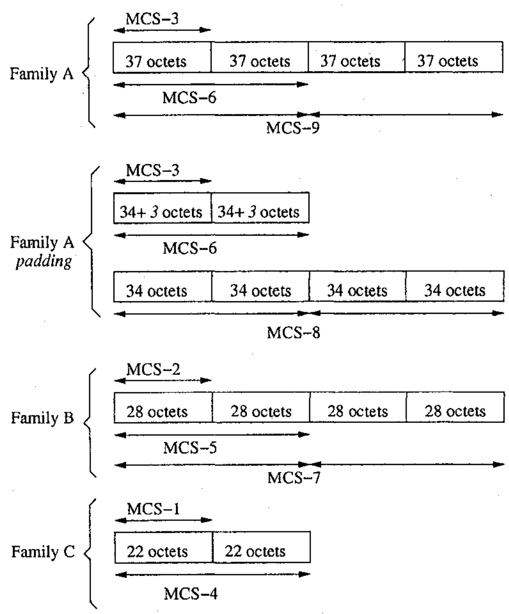


Figure 2.8: Coding Schemes Families

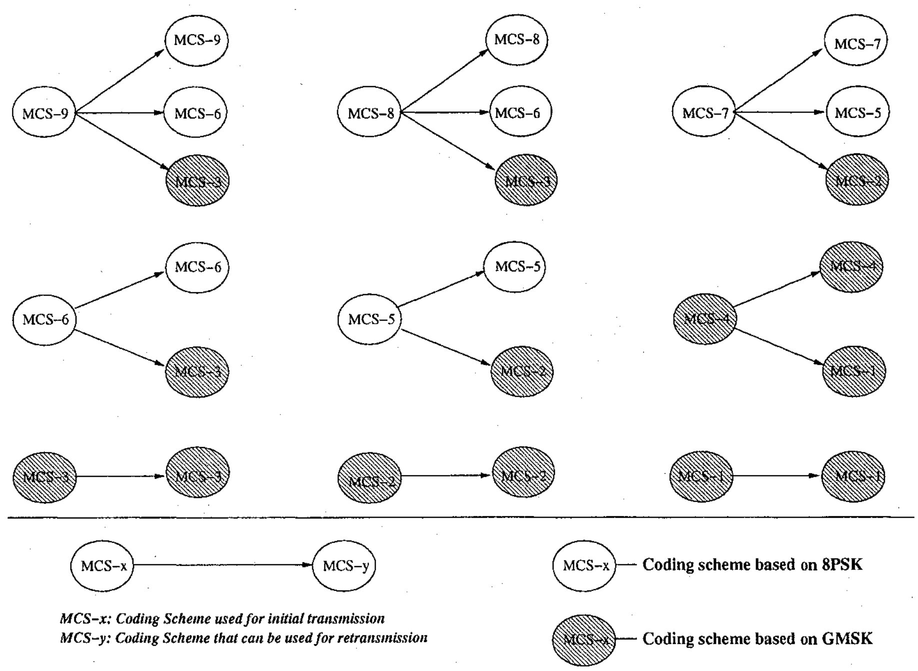


Figure 2.9: Coding scheme chosen for retransmission in EDGE

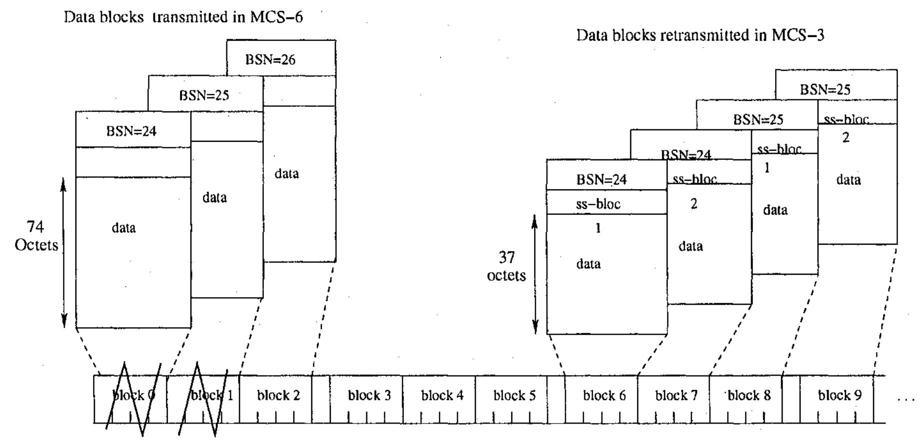


Figure 2.10: Resegmentation in EDGE transmission

The radio blocks are first transmitted in MCS6. The erroneous blocks 24 and 25 need to be retransmitted. And the retransmission is performed in MCS3 due to the worsening of the link quality.

2.3.5 Link Quality Control

Link Quality Control (LQC) is a term used to designate the techniques that allow a choice of a coding scheme according to the current link quality[19]. The purpose of LQC is to maximize link throughput and minimize transmission delay through adapting link robustness to the varying channel quality. The reason of the utilization of the LQC is to improve the efficiency of radio resources.

LQC is necessary for EDGE. As we know, EDGE uses a new modulation named 8PSK to obtain a higher transmission rate. The 8PSK modulation is not as robust as the GMSK modulation, and will not perform equally well in all parts of a cell. Therefore, it is necessary for EDGE to use LQC to adapt the modulation and channel coding to the radio quality of each link.

In EDGE, LQC includes Link Adaptation (LA) and Incremental Redundancy (IR).

2.3.5.1 Link Adaptation (LA)

The principle of link adaptation is the selection of coding schemes according to the varying link quality conditions to achieve a maximum throughput. The channel quality is estimated continuously by the mobile and reported to the network. Then the coding scheme maximizing the throughput is chosen. Figure 2.11 from [12] shows an example of the relationship between throughput and C/I (Carrier-to-Interference ratio, dB) for different coding schemes. Here the interferenece means the co-channel interference. For instance, in the case that C/I is 15dB, the maximum throughput can be obtained when MS7 is used. Therefore, MS7 will be chosen as the current coding scheme.

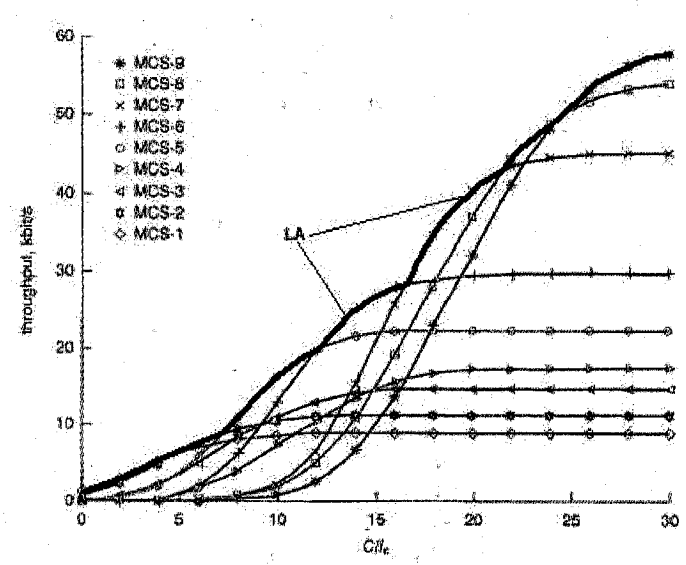


Figure 2.11: EDGE Link Adaptation (cited from [12])

We have seen that EDGE has resegmentation capabilities (see Figure 2.9). In GPRS, however, this resegmentation (retransmitting with another coding scheme) is not possible. Once packets have been sent, they must be retransmitted using the original coding scheme even if the radio environment has changed. It is the main difference between the LQC in GPRS and in EDGE.

2.3.5.2 Incremental Redundancy (IR)

In EDGE, the data block of L bits is first encoded at a rate of 1/3[17], which produces a code of 3L bits. Then the code is punctured depending on the MAC/RLC header of the radio block[19]. That means only a part of the coded bits is picked to transfer through the PDTCH in order to improve the transmission rate. For example, in MCS-9, the block of 612 bits is encoded with the 1/3 convolutional code, which results in a block of 1836 coded bits (C (0),C (1),...,C (1835)). Three puncturing schemes P1, P2 and P3 are applied in such a way that the following coded bits are transmitted:

- P1: {C (3j) for j=0,1,...,611} are transmitted
- P2: {C (3j+1) for j=0,1,...,611} are transmitted
- P3: {C (3j+2) for j=0,1,...,611} are transmitted

Different puncturing schemes are applied for different MCSs.

In IR mode, the RLC block is first transmitted with very little coding (e.g. P1) and the code rate is R_1. Here R_1=uncoded bits/coded bits.For example,in MCS-9, the length of uncoded bits is 612. If puncture scheme P1 is used, only 612 coded bits are transmitted. So $R_1 = 612/612 = 1$. If decoding is successful, a very high throughput can be obtained. If decoding is unsuccessful, more redundant information will be sent by the next retransmission, using a different puncturing scheme. The erroneous blocks are stored (whereas in LA, the error block is discarded) and combined with the information received with the next retransmission. For example, in MCS9, a block is transmitted first with a puncturing scheme P1 ($R_1 = 1$). If the block is received in error, the puncturing scheme P2 will be used in next retransmission. The receiver does not throw the previously received data block, but uses P1+P2 (i.e. 1224 coded bits) to decode the currently received data block, resulting in a lower code rate $R_{1+2} = 612/1224 = 0.5$. This gives the mobile/base a greater chance of correctly decoding the data, which decreases the block error rate (BLER). If all the codewords (different punctured versions of the encoded data block) have been sent, the first codeword (which is punctured with P1) is sent, and so on [18]. In IR mode, the link performance is dependent on the combination of the performance of the packets coded in a particular coding scheme. An example of code combination of IR is shown in Figure 2.12. A radio block of 4 bits is firstly encoded to a 12 bits coded data (c (0),c (1),...,c (11)). For the first transmission, c (0),c (3),c (6) and c (9) are picked to transmitted. If first decoding is unsuccessful, these 4 bits are stored in memory. For the second transmission (first retransmission),c (1),c (4),c (7) and c (10) are transmitted, and the second decoding is based on c (0),c (1),c (3),c (4),c (6),c (7),c (9),c (10). If the second decoding is still unsuccessful, c (2),c (5),c (8)and c (11) will be transmitted for the third transmission (second retransmission). At this time, the decoding is based on c (0),c (1),...,c (11).

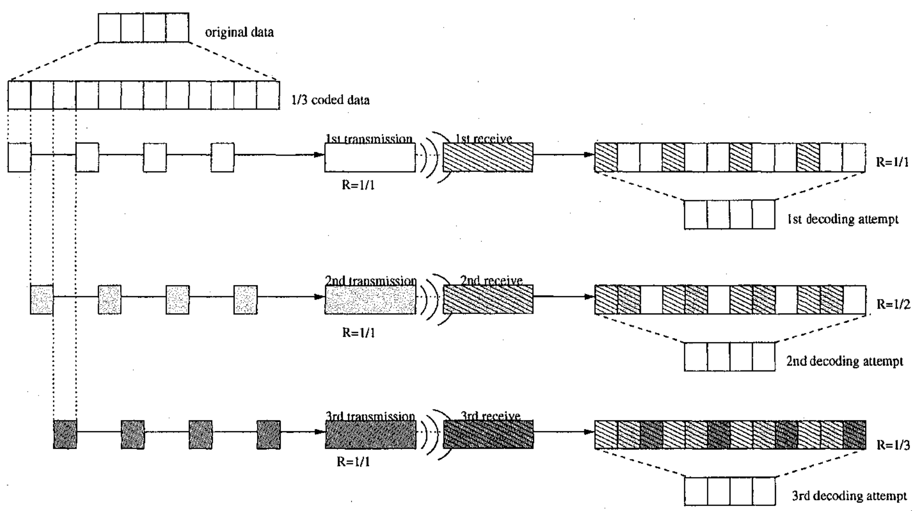


Figure 2.12: Code combination of Incremental Redundancy

2.3.6 Quality of Service

The Quality of Service (QoS) requirements expressed through the QoS profile are defined as follows:precedence, reliability, delay and throughput[28]. A QoS profile is associated with each PDP context.

- The service precedence indicates the relative priority of maintaining the service. For example in case of network congestion which packets may be discarded first. Three priority levels are defined: high, normal and low.
- The reliability indicates the transmission characteristics that are required for an application. There are 3 reliability classes which are defined with different requirements of the following cases[28]:
  - probability of data loss;
  - probability of data delivered out of sequence;
  - probability of duplicate data delivery;
  - probability of corrupted data.
- The delay parameter defines the end-to-end transfer delay incurred in the transmission of service data units (SDUs) through the EGPRS networks.Four classes of delay have been defined in the standard[28].
- Throughput is defined by two negotiable parameters: Maximum bit rate and Mean bit rate. It shall be possible for the network to re-negotiate the throughput parameters at any time during a session.

Table 2.2 gives the QoS parameters defined in ETSI standard[25][28].
EDGE can support multimedia service. The requirements of different applications and users are different. Therefore, more QoS classes must be defined in EDGE. According to the EDGE standard [29], EDGE can support four different QoS classes;

Table 2.2: QoS Profile

QoS parameters	defined parameters
Precedence class	1 (high) - 3 (low)
Reliability class	1 - 5
delay class	1 (short delay) - 4 (Best effort)
Peak throughput class	1 ($\leq 8kbit/s$) - 9 ($> 2048kbit/s$)
mean throughput class	1 - 18, 31 (Best effort)

conversational, streaming, interactive, and background. The main differentiating factor between these classes is how delay sensitive the traffic is: conversational class is very sensitive to delay traffic while background class is the most delay insensitive traffic class.

The *conversational service class* is used for real-time services, such as IP telephony, video-conference, etc. This class needs low transmission delay and low delay variation.

The *streaming service class* is used for one-way real-time service transfer, such as real-time audio and video-stream. It needs constant delay and throughput.

The *interactive service class* is a request-response pattern of data transfer, such as web browsing and telnet. The round-trip delay is an important factor.

The *background service class* is used for best-effort traffic data, such as SMS and email. This kind of service is not sensible to delay.

The end user QoS requirements are shown in table 2.3 (cited from [29]), FER means Frame Error Rate).

To define a QoS contract between the MS and the network, packet data protocol (PDP) contexts containing QoS profiles are negotiated between the MS and the SGSN[27]. The BSS is responsible for resource allocation on a Temporary Block Flow (TBF) base and scheduling of packet data traffic with respect to the relevant

Table 2.3: End-User QoS Requirements

Traffic Class	Medium	Application	Data rate (kbits/s)	delay	Information loss
Conversational	Audio	Telephony	4-25	<150ms	<3% FER
	Video	Videophone	32-384	<150ms	<1% FER
	Data	Telnet,interactive games	<8	<250ms	Zero
Streaming	Audio	speech,music	5-128	<10s	<1% packet loss
	Video	real-time video	20-384	<10s	<2% packet loss
	Data	FTP	-	<10s	Zero
Interactive	Audio	voice messaging	4-13	<1ms	<3% FER
	Data	Web-browsing	-	<4s/page	Zero
Background	Data	Fax,Email	-	>10s	-

QoS profiles negotiated. The tasks of the GGSN comprise mapping of PDP addresses as well as classification of incoming traffic from external networks.

The mechanisms for QoS management in EDGE can be regarded as a three-stage model. On PDP context activation the QoS parameters are negotiated. As long as the PDP context remains active, these parameters should be guaranteed unless there is a QoS renegotiation. The QoS profile is considered both for each TBF and for each radio block period. During the TBF, radio blocks are scheduled at the BSS in competition with other existing TBFs in the radio cell. This scheduling function has to be performed considering the QoS profiles of the PDP contexts associated with the TBFs.

Even though the QoS requirements have been defined in EDGE standard, the QoS guarantee is not defined by the standard bodies. The satisfaction of QoS requirements is left vendor dependent.

2.4 Conclusion

EDGE is an evolution of GPRS and HSCSD and is suitable for circuit- and packet-switched services. EDGE reuses GSM/GPRS core network, GSM carrier bandwidth and time slot structure. The modifications mostly concern the RLC/MAC layer and the physical layer. In order to achieve higher data rates, a modulation scheme named 8PSK is introduced. 8PSK will not replace the current GMSK but coexists with it. Link quality control (LQC) is the core of EDGE, which allows link adaptations as to the varying radio link quality. Without LQC, all links must be designed for the worst case, and evolution of EDGE cannot be represented. In order to simulate user perceived QoS in an EDGE network, the simulation of LQC scheme is necessary.

Chapter 3

Literature Review

In this chapter, we present a discussion on prior literature related to the performance of an EDGE network. There have been several proposals to assess the performance of an EDGE system. Three types of approaches are taken into account in the literature: system level simulation, simulation of protocols and analytical models. System level simulation is realized in a live network and the purpose is to measure the performance of the whole system. In the simulation of protocols, a full or a part of EDGE protocols is implemented in the simulator to enable a realistic study of the behavior of EDGE. As for the analytical model, the performance of a network system can be studied through a mathematical model.

3.1 System level simulation

System level simulations are performed in [2][3][4]. The simulation is realized in a live network. The simulation environment includes a regular cellular layout consisting of a large number of equally sized macro cells. A large number of mobile stations are studied over a period of time. Mobile stations enter and leave the system during the simulation. What the simulation concerns is the system performance, such as system throughput, spectral efficiency, etc. User's QoS is not their purpose. Besides, the simulation is both costly and inflexible since it needs real environment to be built. In such an environment, it is difficult, if not impossible, to change the simulation parameters such as the classes of service, the LQC algorithms, etc.

3.2 Simulation of protocols

The simulation of protocols consists of the simulation of complete protocols and the partial simulation, where a full or a part of the EDGE protocols are implemented in the simulator to estimate the performance of an EDGE network.

3.2.1 Simulation of complete protocols

This kind of simulation can model different protocols of EDGE[5][45][48]. Figure 3.1 depicts the simulation model used in [5] and [45].

This simulator is based on the programming language C++ which simulates the exchanges between the different layers of (E)GPRS. The simulator comprises the modules MS, BS, SGSN, the transmission links, the load generators, a user interface for presentation and a module for statistical evaluation. The respective protocol stacks are implemented in MS, BS and SGSN. The module *channel/error model* defines the state of the radio channel and the mapping between the actual *Carrier-to-Interference* (C/I) and the *Block Error Probability* (BLEP). The module *Channel Management* is responsible for the channel allocation. The detailed implementation of the standardized protocols enables a realistic study of the behavior of (E)GPRS. As we can see, the simulator models precisely each protocol layer as well as the exchanges between these layers. However, the simulation is hard to develop and lacks flexibility. In addition, the results of these works only consider the system performance, such as system throughput, channel utilization, mean packet delay, etc. Our objective is to produce the results at message and user level in order to evaluate the quality of service offered to the user. We could try to make a precise

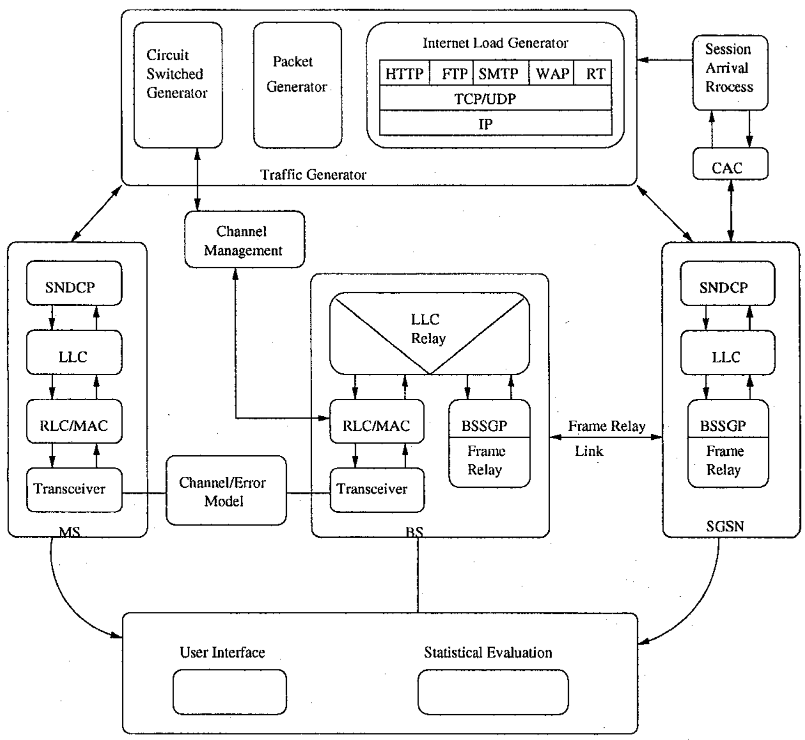


Figure 3.1: Structure of the (E)GPRS simulator[5][45]

simulator to achieve our goals, but the complexity of development as well as the lack of flexibility make that we prefer to draw aside this possibility.

3.2.2 Partial Simulation

In the literature, we can also find the partial simulation [12]. In this simulation, only a part of the network is considered, and the results are LLC frame delay and throughput. Many of the simulations focus on a particular point of the protocol, for example: link layer performance; MAC layer performance (packet scheduling and resource allocation), etc.

3.2.2.1 Simulation of link layer performance

Some of the simulations estimate the link layer performance[44][46][11][39][10][37][50][6]. Different LQC (Link Quality Control) algorithms are studied and compared. It would be a relative fast way to get the message of the performance of the different lQC algorithms in an EDGE network. But the results only concern the system throughput by means of channel condition and resource utilization.

3.2.2.2 MAC layer performance

We can also cite [8][52][40][14][34] and [36], where different scheduling and resource allocation strategies are considered.

[40] tested the behavior of channel allocation mechanisms under different conditions. It only concerned the frequency efficiency under different channel allocation. User perceived QoS is not its objective.

In [34] and [36], packet scheduling and radio resource management are investigated for providing differentiated services in EGPRS system. Two classes of background data service are considered. In order to differentiate between premium and basic users, the scheduler uses a weighted round-robin scheme. The performance measures are *file transmission rate* and *normalized file delay.* The former is defined as the ratio of the sum of all file sizes and the sum of all file transmission delay, whereas the latter is defined as the ratio of the file transmission delay to the file size. But the transmission delay considered in the simulations is just the time required to deliver a file from the SGSN to the MS. What we want is an end-to-end result, i.e. transmission time from packet network to user.

In the articles [8][52] and [14], QoS gains provided by the different scheduling algorithms are shown. Let us see if they can fulfill our objective. The user's QoS requirement used in [8] is normalized delay (s/kbit), which is defined as total absolute delay (queuing time plus transmission time) divided by packet size in kbit. Packet data that cannot be immediately transmitted are queued and sent as soon as resources become available. The objective of the simulation is to find a scheduling algorithm that can maximize a fraction of satisfied users. Packet discarding rate is not considered in this simulation even though users in a bad radio link situation will be dropped. No voice calls are considered either. For simplicity, all users have the same QoS requirements. Besides, one user can only occupy one channel for transmission, so that channel multiplexing cannot be represented in this simulation. In [52], the QoS parameter is the average throughput per user. It can be used to compare different scheduling algorithms, but it does not accurately describe the user's effective experienced delay. [14] built a computer based model to enable the testing of resource allocation algorithms in data transmission over EGPRS networks. The results are the total packet delay. But for simplicity, the users can only use coding scheme one and two so that the flexibility is relative low. Moreover, the influence of the voice transmission is not considered in the simulation. In a word, the purpose of these simulations is the comparison of different scheduling algorithms but not the user's QoS in an EDGE network.

3.3 Analytical models

In addition to the simulations mentioned above, an analytical model can be a faster way to study the performance of a network system and can give satisfying results. For example, [38] provide a new medium access method to avoid the instability of Slotted Aloha mechanism. [51] investigate the uplink performance over the LLC frame level, which is more useful in network design and optimization. Whereas in [42],[43] and [41] an analytical model is used for the performance of downlink transmission. F.Bodevin also provided some analytical models in [31] for both uplink and downlink transmission.

In the following subsections, we analyze these analytical models to see what they can give us. We try to keep in mind the objective of our research, that is, to develop a tool to evaluate the user perceived QoS. Therefore, only models of this direction will be studied.

3.3.1 The models for uplink transmission

The radio channel requirement for the uplink transmission is based on the Slotted Aloha mechanism. And the model is more complicated than that for downlink transmission.

In [38], the system is modelled as a state-transition diagram (see Figure 3.2). P_{ij} is the probability of going from state i to state j for $i, j = 0, 1, 2$. The arrival process is assumed to be Poisson. The throughput and the blocking probability can be obtained from the steady-state probability formulas of the Markov chain. The throughput reaches the maximum value when traffic loads reaches a value G_{max}. For traffic loads higher than G_{max}, the throughput decreases drastically. In order to avoid the occurrence of this situation, for normalized traffic loads higher than G_{max} all users in the system are assigned a probability of transmission G_{max}/G and are not allowed to transmit with probability $1 - G_{max}/G$. In this way, the throughput

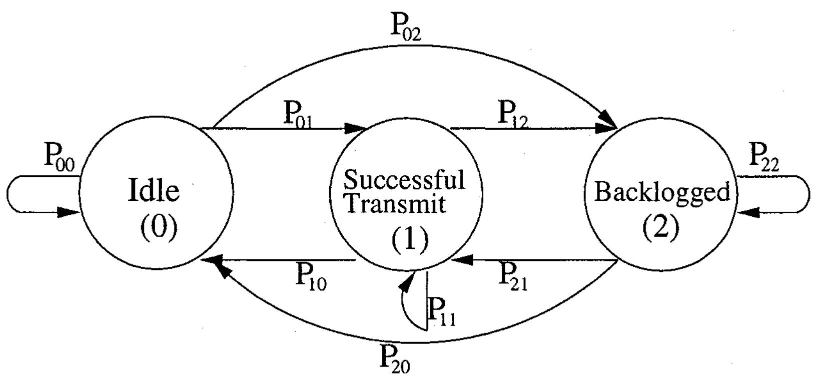


Figure 3.2: Markov chain for the mathematical model (cited from [38])

will not decrease significantly when traffic loads exceed G_{max}.

This model can be used to analyze the throughput and blocking probability, but the transfer delay cannot be obtained through this model. Besides, this approach is only valid for the case that the uplink transmissions take place over one timeslot. We can remedy this problem by using a model proposed in [31].

In [31], The behavior of slotted Aloha is described as a discrete-time Markov chain (see Figure 3.3). There are always N mobiles that are liable to transmit the requests. The steady-state probability formulas can be used to solve this Markov chain, and the mean throughput and the mean waiting time can also be deduced.

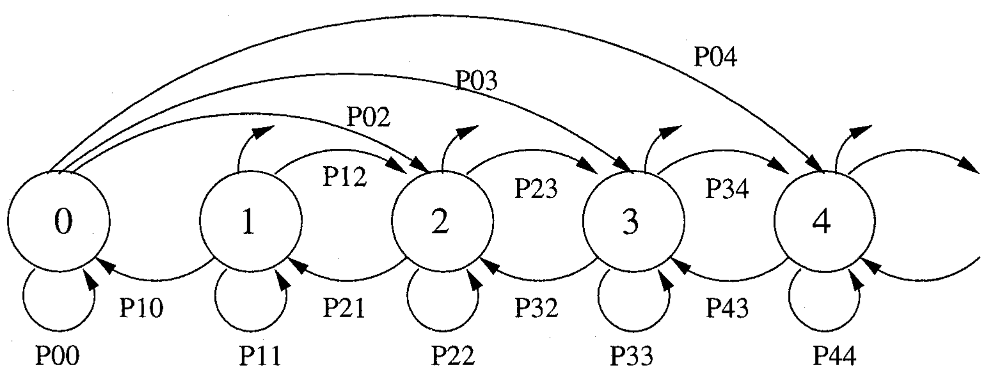


Figure 3.3: Markov chain for Slotted Aloha (cited from [31])

In fact, according to [9], this system is not stable and can vary between two stable points. The departure rate is almost zero whenever the system jumps to the undesirable stable point. This is the characteristic of this model. In this case, it will be necessary to check the validity of the mean values obtained from the Markov chain. In the next subsection, we will study the models used for downlink transmission.

3.3.2 The models for downlink transmission

There are several queueing models proposed in the literature, such as the closed queueing model[42][43], the model with buffer and the circuit switching model without buffer[31].

3.3.2.1 Closed queueing network model

In [42] and [43], a simple queueing network model is developed to estimate the user perceived performance of interactive data application. To characterize the user's experience, they use end-to-end page delays and introduce a measure called Equivalent Circuit Rate (ECR). The Equivalent Circuit Rate for a user of a shared packet access network is the circuit rate required by an identical user with a dedicated connection in order to achieve the same user-perceived performance (e.g. page delay). The system can be modelled with the closed queueing network model shown in Figure 3.4([42][43]).

The figure shows M interactive users with ON/OFF traffic, who are sharing an access network queue. A mean latency E[D] is used to represent the protocol overhead and backbone propagation delays. The total web page delay time comprises the service time, the waiting time and the overhead delay. Assume that both the OFF time and the service time are exponentially distributed. The Equivalent Circuit Rate (average delay time) can be deduced from the steady state diagram of the closed queue.

The problem of this model is the exponential distribution of the service time. In fact, the service time depends on the web page size and on the channel rate. In

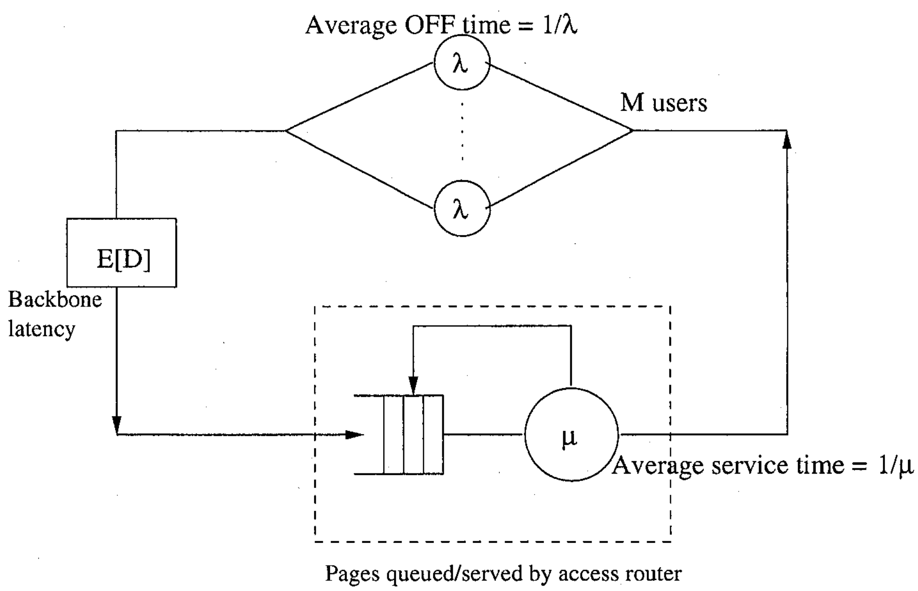


Figure 3.4: Queueing model for M interactive ON/OFF sources sharing an access network (cited from [42][43])

this model, we cannot choose the distribution of the web page size. Besides, we cannot provide the different quality of service to the different classes of service. The result will be useful in network dimensioning and admission control, but it is hard to evaluate the QoS offered to user in an EDGE network.

3.3.2.2 Model with buffer

M/M/C/N model

Assume that the arrival of data users is a Poisson process with rate λ and the service time is exponentially distributed with a mean of μ. For the transmission of the single slot service in a fixed number of C channels, the average queueing time can be obtained from the M/M/C/N queueing system, where N is the maximum number of users in system (in service and in queue)[31]. Using this queueing model, analytical results can be easily obtained.

However, this approach does not permit to differentiate the quality of service according to the priority of the users. The M/M/C/N model with management of queue is one of the solution of the problem[31].

M/M/C/N model with management of queue

Two classes of service A and B are considered in this model. Every arriving packet of each class has a probability of being accepted by the system. The blocking probability, the throughput and the mean waiting time of each class can be deduced with the help of steady-state probability formulas [31].

The problem of this model is the exponential distribution of service time. We can neither choose the distribution of the messages, nor give the priority to user groups according to the message size.

M/G/C/N system is supposed to be a more realistic model, where the exponential process is replaced by a general distribution. However, this system is more complicated to analyze. Even though there exists some formulas for approximating the average waiting time for the M/G/C systems, we still need a numerical simulation to obtain the results of the parameters of the QoS in a M/G/C/N system. There is no analytical models for such a system.

3.3.2.3 Circuit switching models without buffer

Independent channels model

The circuit switching model is chosen because the succession of assignment is not interrupted and the radio block of the same message are sent with the same quality of service, which is equivalent to a circuit switching. In the circuit switching model without buffer, the transaction is rejected if the call is not accepted (see Figure 3.5 from [31]). In this case, the only QoS parameter is the blocking rate.

The approximation of the blocking probability can be obtained from [31].

However, the approach permits to analyze the blocking rate for one frequency band but not for several bands. If one wants to analyze several frequency bands, the problem of fragmentation of free resources will appear [31]. In order to resolve the

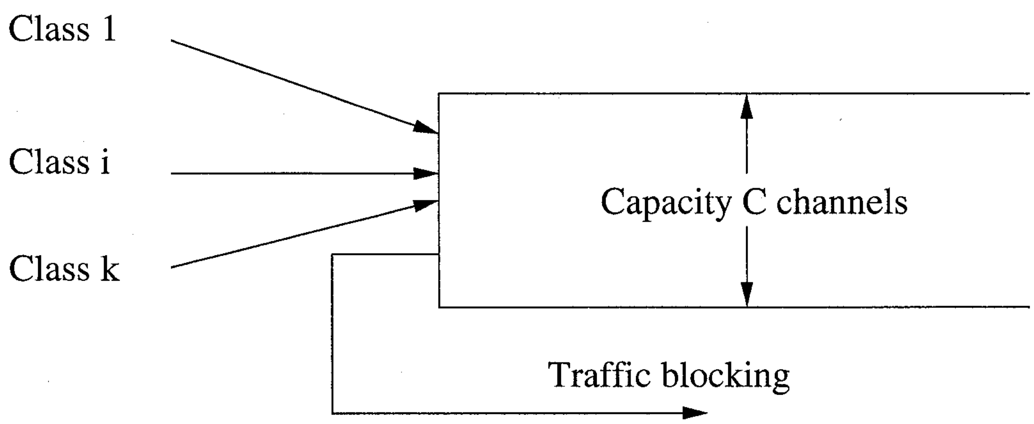


Figure 3.5: Circuit switching model without buffer (cited from [31])

problem, the possible solutions are proposed in [31].

Independent bands model

Let us consider the system of Figure 3.6. Each frequency band is distinct, which eliminates the problem of fragmentation. Let m distinct frequency bands, each one having C/m (E)GPRS channels, and k classes of traffic.

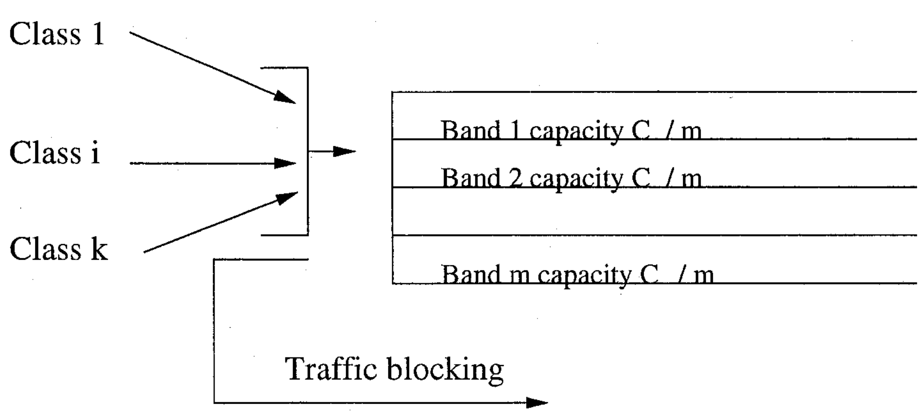


Figure 3.6: Model with distinct frequency bands (cited from [31])

There are 2 possibilities for this model [31]:

- **Overflow not allowed.** The calls are randomly directed, with a uniform probability, to one of the m bands and will be dropped if this band cannot accept it (see Figure 3.7).

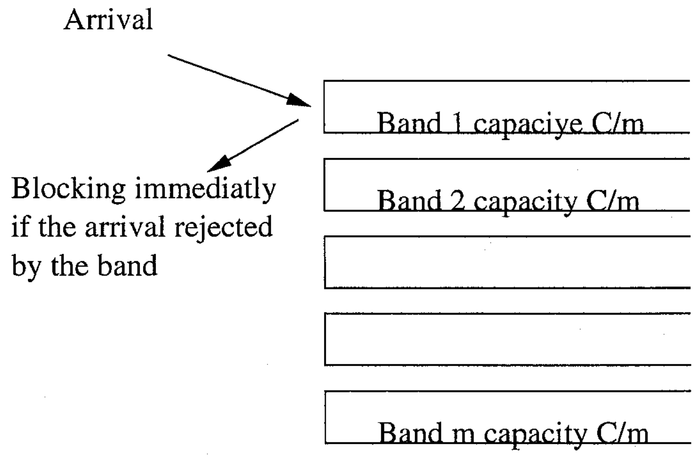


Figure 3.7: Distinct frequency bands without overflow (cited from [31])

- **Overflow allowed.** If there is a place in one of these bands, then the call will be accepted (see Figure 3.8).

In the first case, the system comes down to identical simple models with k classes of arrival rate λ_i/m for each band. In this case, it has the same blocking probability as a system with 1 band, of capacity C/m with k classes of arrival rate λ_i/m, of times of service μ_i and which require C_i channels.

In the second case, the system cannot come down to simple model any more and it is necessary to find the state distribution of the Markov chain. However, there is a simple model for this case (see Figure 3.9 [31]). Here, the bands are traversed in a sequential way. If there is no place in the first band, the blocked call is sent to the second one, etc.

The model without overflow and the model with overflow can be mixed into a mixed model in case that the Poisson approximation is not valid any more for the arrivals

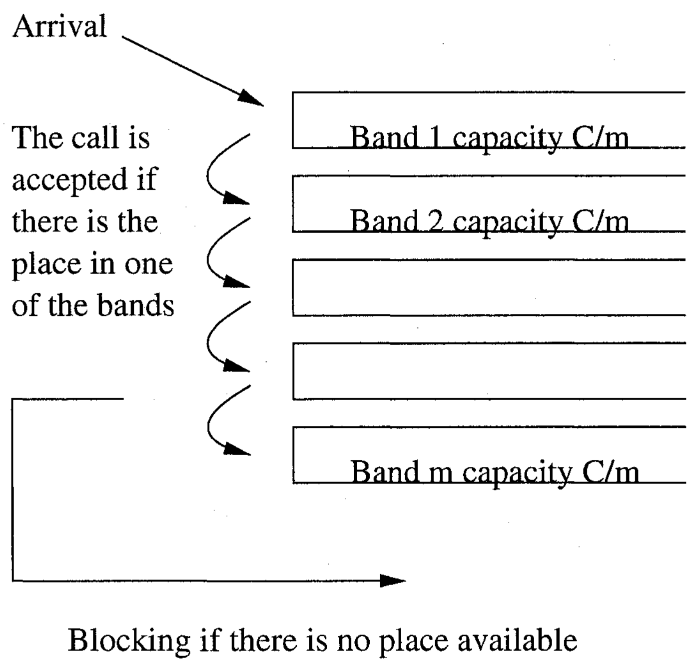


Figure 3.8: Distinct frequency bands with overflow (cited from [31])

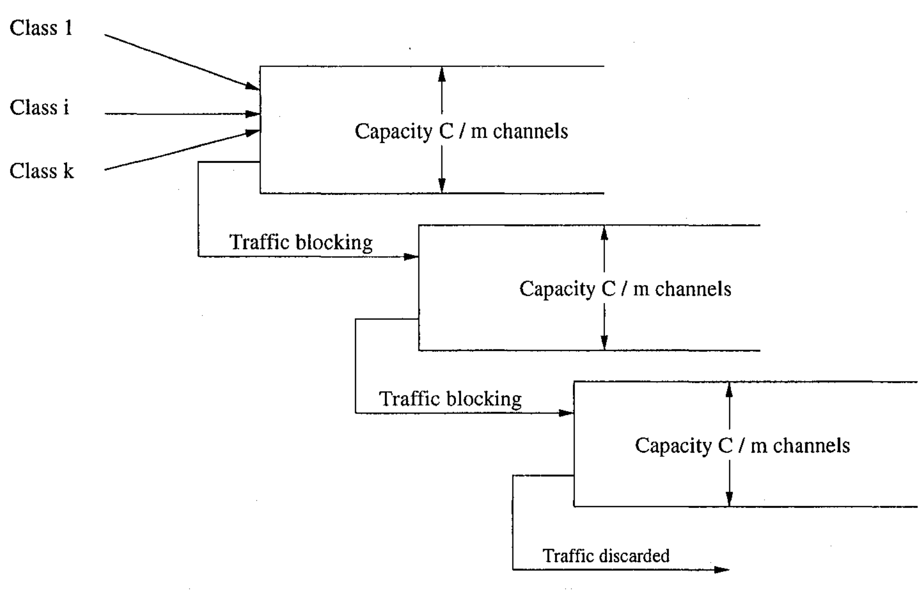


Figure 3.9: Model with separate frequency band and ordered path of the bands (cited from [31])

to bands m, $m-1$,... One can consider a mixed model with the sequential course as long as the flow of blocked calls can be considered as Poisson and then the model of the independent bands for the remainder of the frequency bands (see Figure 3.10).

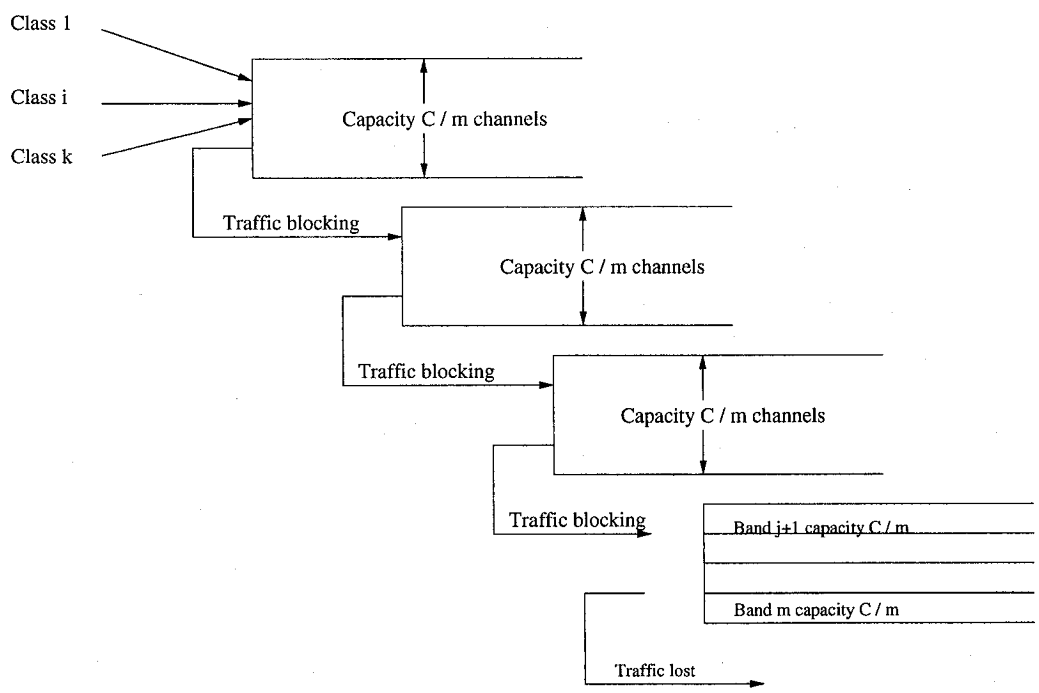


Figure 3.10: Mixed model (cited from [31])

To recapitulate the various models considered, let us consider a cell with C channels on m bands, k classes of traffic of arrival rate λ_i, of service time μ_i and which need C_i channels. Let $P_b^i(C, \lambda_i)$ the blocking rate the class i in the system(see Figure 3.5). Then the blocking probability can be found in Table 3.1 [31].

3.3.2.4 Independent channel model with buffer

The models mentioned above do not take into account a possible waiting time before the allocation of resources for transmitting. First of all, there is the time inherent in the signalling protocols, for example, to inform the mobile that it will receive a data transmission on such channels. This takes short time, maximum several frames. We

Table 3.1: Blocking probability following the used model

model	Classial	Without overflow	With overflow	Mixed
Blocking	$P_b^i(C, \lambda_i)$	$P_b^i(\frac{C}{m}, \frac{\lambda_i}{m})$	$P_b^i(\frac{C}{m}, \lambda_i P_b^{i1} \ldots P_b^{i(m-1)})$	$P_b^i(\frac{C}{m}, \frac{\lambda_i P_b^{i1} \ldots P_b^{ij}}{m-j})$

can also put the blocked calls into the buffer. In the previous models, the arrival calls will be discarded if there is no free resources. However, as in the case of M/M/C/N model, the blocked calls are not discarded automatically but are placed in a buffer before being able to be served. We can integrate a buffer into the precedent models, for example, the model with overflow (see Figure 3.8). Then we have the system with overflow and buffer (see Figure 3.11).

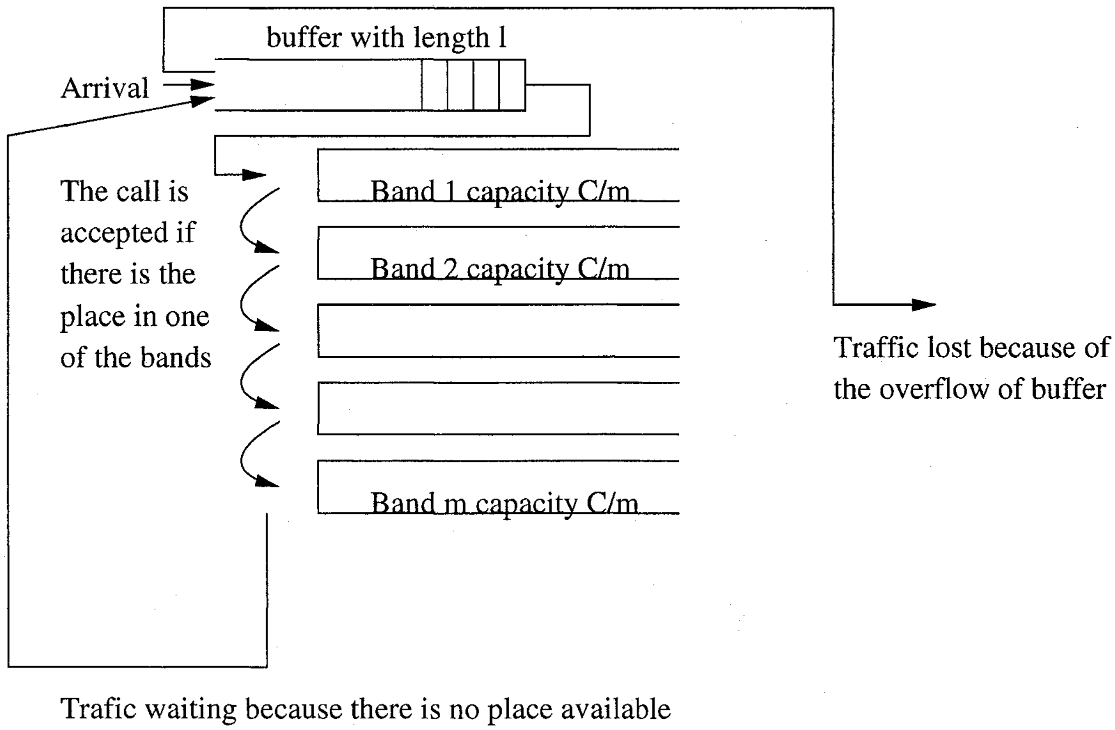


Figure 3.11: Distinct frequency bands with overflow and buffer (cited from [31])

The modifications are very few for the model with overflow because it did not put the communications into buffer until there is no any place to accept it. For the model without overflow, the modifications are significant because it is possible to putting in buffer in each case that the arrival calls are rejected by certain channel.

As we have seen, none of these models is ideal. The queueing models only consider one channel per communication and one cannot introduce classes of service according to the size of the messages, or use a Weighted Fair Queue (WFQ). The circuit switching model without buffer makes it possible to manage the multi-channels communications as well as several frequency bands but do not take into account the delay. The circuit switching model with buffer makes it possible to measure the delay, blocking and allows the communications on several channels, but the classes of service are always limited. For example, one cannot introduce classes of service according to the message size, or use a Weighted Fair Queue (WFQ). Besides, these models do not take into account the impact of the voice call, which is very important in an (E)GPRS network. And in order to obtain the theoretical result, the choice of statistical distributions of arrival of the messages and service time are also limited.

In brief, the above work cannot be used to evaluate user perceived Quality of Service in an EDGE network. Therefore, we decide to propose our simulation model that we will describe in the next chapter. The model is based on the GPRS simulator proposed in [31] and [32].

Chapter 4

Simulation Model and Implementation

In this chapter, a simulation model will be provided to evaluate the quality of service offered to a user in an EDGE network. But before doing that, we review the presentation of the simulator proposed in [31] and [32].

4.1 GPRS simulator

This simulator is used to evaluate the user perceived Quality of Service in a GPRS network. Figure 4.1 depicts the simulation model.

First of all, the data messages produced by the traffic generator are assigned to the different service groups according to the different QoS policies. After the radio resources (timeslots) used for data transmission are decided, the messages will be sliced into RLC blocks and will be sent to certain queues. Then, the logical channels will be allocated to the users and the transmission turns will be assigned to each class of service to begin the data transfer.

The traffic includes voice calls and data messages and the message size is randomly

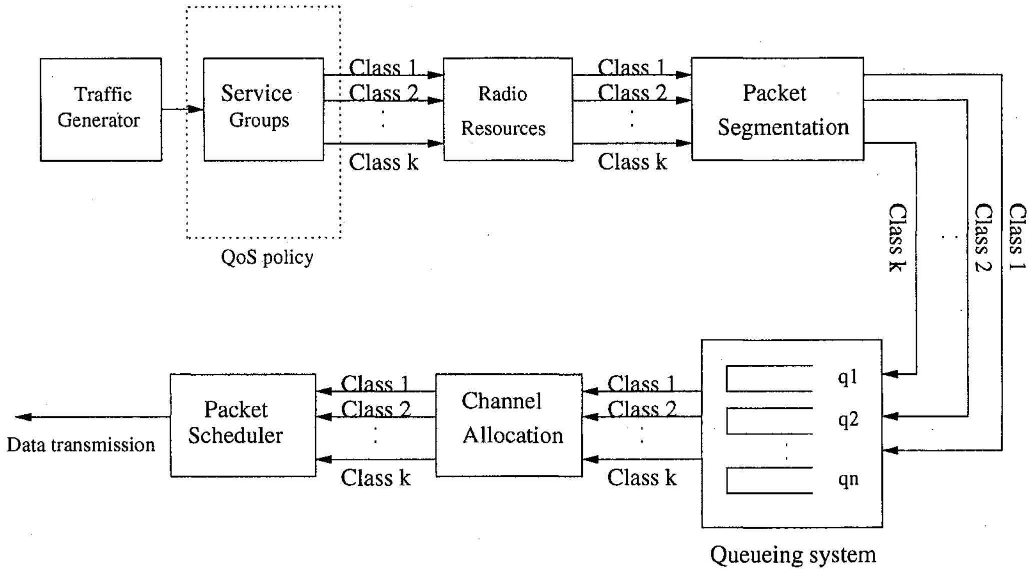


Figure 4.1: GPRS Simulation Model

distributed. The arrival packets can be classified into different service groups according to the current QoS policy and the following QoS policies are proposed:

- QoS policy based on the message size.
- QoS policy based on privileged users.

Different weights are assigned to different service groups.

The GPRS simulator assumes one radio frequency per cell (8 timeslots). One timeslot is used for signalling, up to 4 timeslots can be shared between voice and data or reserved only for data, the remaining 3 timeslots are reserved for voice only. Voice calls have a preemptive priority (see Figure 4.2[32]).

The allocation of resources (channel allocation) is considered at the block level, that is, the RLC protocol. The messages are sliced into fixed length radio blocks (see Figure 4.3). The allocation is carried out at superframe level. A superframe makes up the 12 radio blocks B0,B1,B2,...,B11. One or several logical channels can be defined for each superframe (see Figure 4.4[32]). At the beginning of each superframe, one

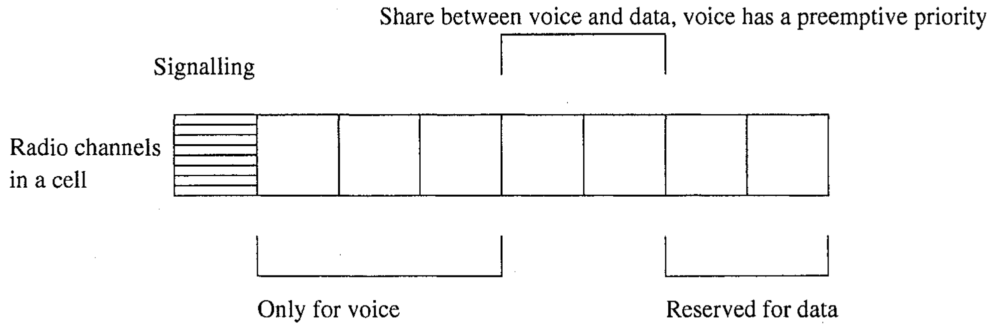


Figure 4.2: Share of radio resources between voice and data (cited from [32])

or several logical channels are allocated to each message. Thus, several different messages can occupy the same physical channel or one message can occupy several physical channels.

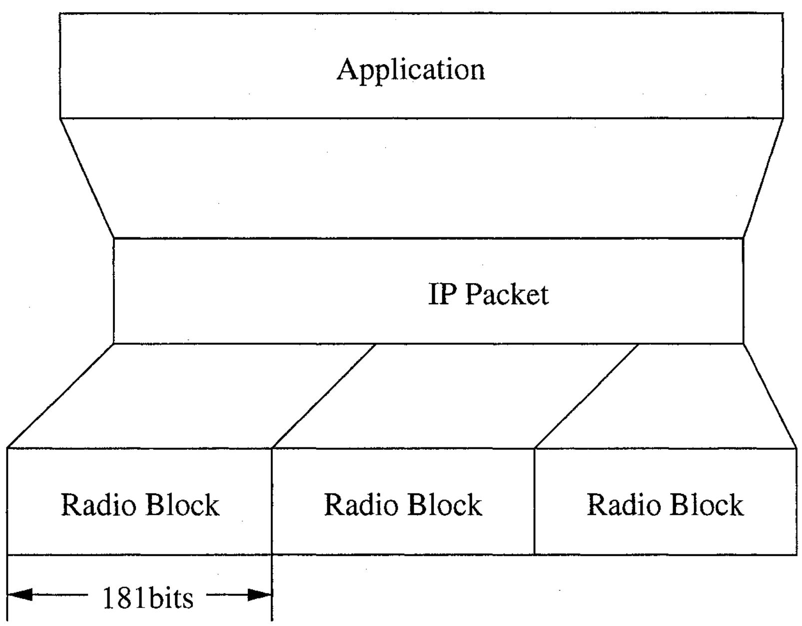


Figure 4.3: Message segmentation (cited from [32])

Once the resources, timeslots have been assigned, the users can be multiplexed with other users sharing the same channels, which is the functionality of MAC protocol[19]. Multiplexing is controlled by the scheduling mechanism. The packet

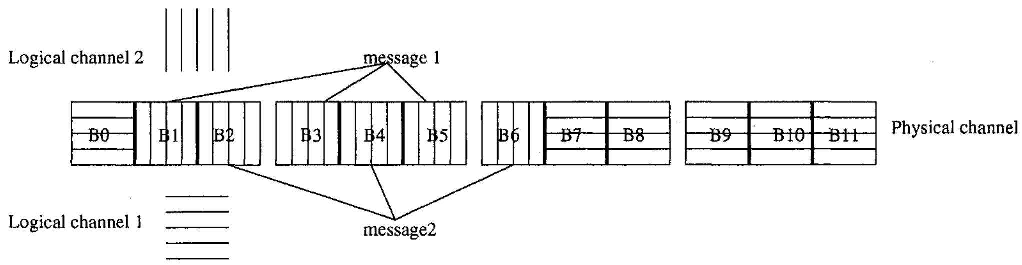


Figure 4.4: Logical channels (cited from [32])

scheduler assigns turns to each user depending on certain criteria. The quality of service is assessed at the message level and the measurement of user QoS are *the delivery time*, *the delivery time for the first block of data* and *the blocking.* As we have seen, in this GPRS simulator, the RLC protocol, classes of services and resource allocation can be modelled. The output measures of user QoS can reflect the total time a user have to wait to transfer his/her message, the system response time and the availability of the system. All of these can be used to estimate a user perceived QoS in an EDGE system. However, in order to give an accurate model to simulate an EDGE system, several modifications need to be made.

4.2 EDGE simulator

In the GPRS model mentioned above, the technology of link quality control is not taken into account, that is, the coding scheme is unchanged and a fixed block error rate (BLER) is used during the process of the simulation. But in an EDGE simulator, a fixed coding scheme can not represent the evolution introduced by EDGE since there is not a single optimal physical layer parameters setting for the varying channel conditions. EDGE introduces a new modulation named 8PSK to obtain a higher transmission rate. The 8PSK modulation is not as robust as the GMSK modulation used by GPRS, and will not perform equally well in all parts of a cell. Therefore, it is necessary for EDGE to use *Link Quality Control* (LQC) to adapt the modulation and channel coding to the radio quality of each link. Link quality control is also the core of an EDGE network. Therefore, an EDGE simulator must contain the functionality of link quality control.

The EDGE simulation model is shown in Figure 4.5.

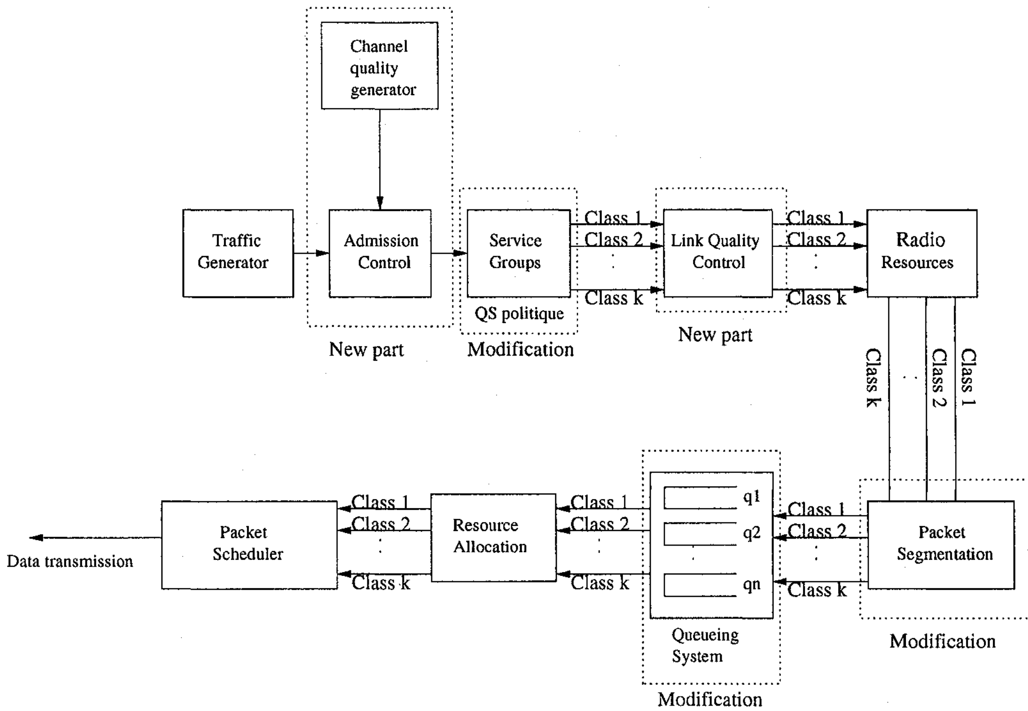


Figure 4.5: EDGE Simulation Model (IP packet transmission)

Once a message is produced, its channel condition will be studied so that the admission control can decide if this message will be accepted or not. If it is accepted by the system, the message will be assigned to one of the service groups according to the different QoS policies and the coding scheme used for the current transmission will be chosen by link quality control. After assigning the radio resources(timeslots) used for data transmission, the message is sliced into RLC blocks and is sent to a certain queue. And then, logical channels and transmission turns will be allocated to this user to begin the data transmission.

The traffic generator and radio resources are similar to those of the GPRS simulator described above.

Then we want to talk about the modules different to GPRS simulator.

4.2.1 Channel quality generator

As we have mentioned above, the link quality control is based on the measurement of the current link conditions. The channel quality estimation is complicated. Many situations have to be considered, e.g. frequency reuse pattern, cell radius, the position of cell phones, etc. Unlike the simulator used in [5] where all of these parameters are taken into account, we just use a channel quality generator to produce a parameter to describe current channel condition.

This parameter could be the block error rate (BLER) and the carrier-to-interference ratio (C/I). Here the BLER is the ratio of blocks received in error to total number of received blocks. In our simulator, carrier-to-interference ratio (C/I) will be chosen as the parameter of current channel quality, where the I means the co-channel interference. The reason is that the distribution of C/I can be seen as log-normal in downlink direction[49], whereas the BLER is randomly distributed between 0 and 1 (uniform distribution) and hard to represent in our simulator. Even though we can use a random generator to produce a value of BLER of current coding scheme, the BLER of other coding schemes should also be calculated to make a decision of coding scheme. Therefore, we still need C/I to calculate these values.

Given a certain C/I, the BLER of a given coding scheme, $BLER_{MCSx}$, can be calculated as follows:

$$BLER_{MCSx} = 1 - (1 - BER(a1, a2))^N \tag{4.1}$$

where BER(a1,a2) is the bit error rate of a given coding scheme at a given C/I; N is the RLC block size in bit of a given coding scheme (e.g. N=592 for MCS6). The

BER can be calculated using an error function, which has parameters (a1 and a2) that represent the effects of modulation and different code rate. The parameters a1 and a2 can be obtained from the data and the curves mentioned in GSM 5.05[23] and 3GPP Tdoc[1] (see Eq.4.2).

$$BER(a1, a2) = 0.5 * erfc(a1 * sqrt(C/I) + a2) \tag{4.2}$$

4.2.2 Admission Control

The purpose of the admission control function is to accept or refuse new users in the radio access network. The admission control tries to avoid overload situations basing its decisions on interference and resource availability measurements.

In our simulator, the admission control is based on the radio link quality measurement. In fact, the transmission with bad link conditions would result in long delays, further degrading the system performance. Therefore, the refusal of transmission with bad link conditions would relatively improve the quality of service offered to users.

Once a packet is produced, a random variable with log-normal distribution is generated to describe the current link quality. If this variable is lower than certain value I (e.g. $5dB$), which means bad link quality, the transmissions to the user will not be permitted and the packet will be dropped. Given the different value of I, we can compare the user QoS to users with and without admission control.

4.2.3 QoS policies

We assume that the arrival packet data will be classified into service groups based on the following QoS policies:

- **Best effort policy.** With best effort policy, all the users have the same priority and the radio resources will be equally assigned to each user.

- **QoS policy based on the message size.** The transmission of large messages will occupy much of the network resources. To avoid exhausting the resources with large messages, the messages are grouped based on their sizes. Then a QoS policy would be modelled by a Weighted Round Robin (WRR) system that will give priority to short messages over large ones. For example: Short messages can be sent to group1, middle messages to group2 and large messages to group3. Then we can give group1 a weight of 3, group2 a weight of 2 and group3 a weight of 1. Therefore, the users in group1 can get more transmission turns and can finish their transmission more quickly. On the other hand, the users in group3 only receive the least transmission turns so that the system will not be monopolized by the transmission of large messages.

- **QoS policy based on differentiated users.** When using this QoS policy, the users are grouped based on their traffic classes without thinking about the message sizes. In this simulator, we use the traffic classes defined in EDGE standard [29][13]. For the time being, only non-realtime interactive and background traffic classes will be considered. Therefore, these traffic classes are:*interactive 1*, *interactive 2*, *interactive 3* and *background*. Different weights are assigned to each traffic class. For example, the users belongs to *interactive 1* will be sent to group1, the users of *interactive 2* to group2, the users of *interactive 3* to group3 and the users of *backgroud* to group4. Group1 will be given the highest weight, whereas group4 will be given the lowest weight. Therefore, more transmission turns will be assigned to group1 in order to guarantee the QoS requirements of these users. The queueing system used for differentiated users will be described in section 4.2.7.

4.2.4 Link Quality Control

When an IP packet arrives, the value of the C/I is studied. Then, the block error rate (BLER) of each coding scheme will be calculated. Depending on the algorithm of link quality control that one wants to use, a different modulation and coding scheme will be decided. According to the EDGE standard[18], there are two Link Quality Control schemes in an EDGE network: link adaptation (LA) and incremental redundancy (IR). All EDGE capable MS can support LA and IR reception. However, which LQC scheme will be used in an EDGE network depends on the system operator. One can use either of the schemes or the combination of these two schemes. In our simulator, the two following LQC algorithms will be taken into account:

- A pure link adaptation (LA) scheme
- A combined LA (link adaptation) and IR (Incremental Redundancy) scheme

Note that the pure IR scheme is not taken into account because when a pure IR scheme is used, the coding scheme will not change no matter how the channel quality changes. In that case if this coding scheme is less protective (e.g. MCS9), a bad channel condition will result in large numbers of retransmission, further degrading the system performance. On the other hand, if this coding scheme is highly protective (e.g. MCS1), the high bit rate will not be achieved.

4.2.4.1 A pure link adaptation scheme

The principle of link adaption is the transmission of code word according to the channel conditions[18]. However, the decision on the coding scheme to use based on the link quality is not defined in the standard. Normally, the coding scheme maximizing the throughput is chosen to improve the system performance. In our simulator, we want to choose a coding scheme that minimizes the transfer delay. Taken the same channel condition (the same C/I), the value of the block error rate (BLER) depends on the current channel coding.

The block error rate will results in the retransmission of error blocks. In our simulator, the retransmission mechanism is represented by simulating an increase in the number of radio blocks.

Assume an IP packet with length L. We can calculate the number N_b of radio blocks that are used to transmit the packet:

$$N_b = \frac{L}{T * R_{MCSx}} \tag{4.3}$$

where T is the mean transmission time per radio block ($T = 20ms$) and R_{MCSx} is the data rate of a given coding scheme (see table 2.1). For instance, assume that current coding scheme is MCS6. Then the raw data contained in one radio block (20ms) is 592 bits, which means the data rate R_{MCS6} of MCS6 is 29.6 $kbits/s$. Let $BLER_{MCSx}$ be the block error rate of a given coding scheme at a given C/I without code combining, we assume that the channel condition will not change during the transmission of the whole packet. The transmission errors will result in using more $BLER_{MCSx}$ of the radio blocks for each transmission and retransmission. That is: $N_b * BLER_{MCSx}$ radio blocks need to be used for the first retransmission, $N_b * BLER_{MCSx} * BLER_{MCSx}$ for the second retransmission, $N_b * BLER^3_{MCSx}$ for the third retransmission, etc. Therefore, the total number of radio blocks N_{total} that has to be used for transmission can be described by Eq.4.4 and the mean transfer delay T_{delay} will be calculated as Eq.4.5:

$$\begin{aligned} N_{total} &= N_b + N_b * BLER_{MCSx} + N_b * BLER^2_{MCSx} + \dots \\ &= N_b * \sum_{n=0}^{\infty} BLER^n_{MCSx} = \frac{N_b}{1 - BLER_{MCSx}} \qquad (4.4) \\ T_{delay} &= N_{total} \times T = \frac{L}{R_{MCSx} * (1 - BLER_{MCSx})} \qquad (4.5) \end{aligned}$$

In our simulator, the decision of the channel coding is to minimize the transfer delay described in Eq.4.5 by considering not to excessively increase the number of retransmissions, since large retransmissions will cause system instability. We assume that the maximum BLER that can be tolerated by the system is 20%. The modulation and coding scheme decision based on the BLER and transfer delay is shown in Figure 4.6. In this figure, T_n represents the transfer delay when using MCSn, and $BLER_n$ the block error rate of a given MCSn.

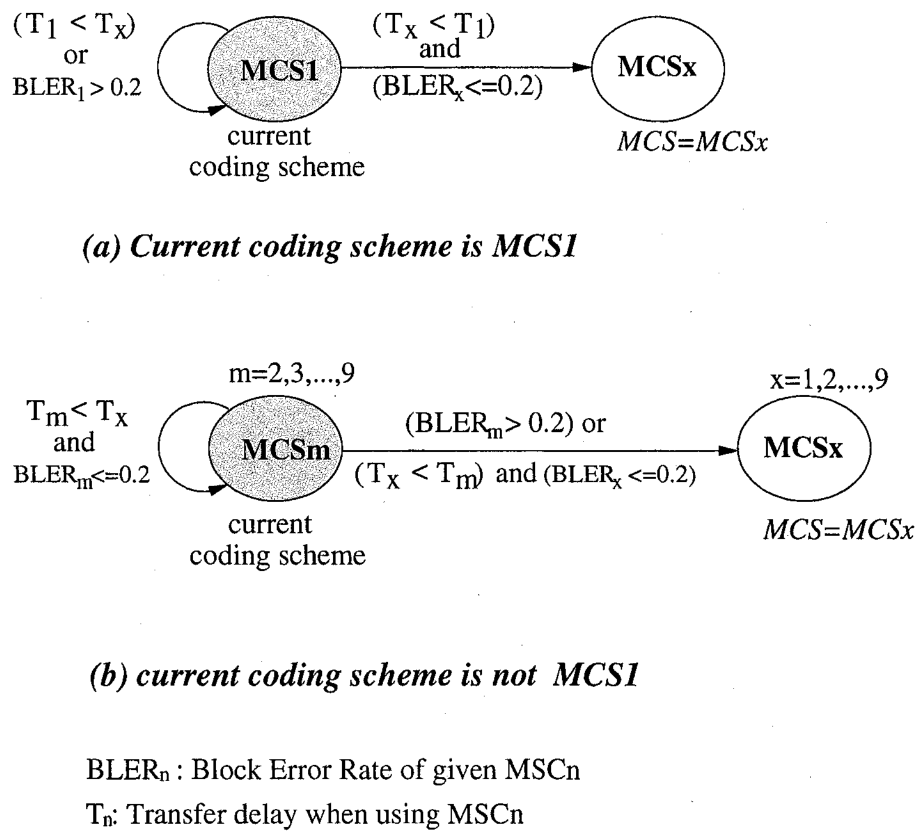


Figure 4.6: Coding Scheme Update and the Corresponding BLER Threshold

Let MCS represent the coding scheme chosen by link adaptation and k be the number of the coding scheme. The procedure for the update of the coding scheme can be explained as follows:

- **The current coding scheme is MCS1**

  The coding scheme used for transmission will still be MCS1 if $BLER_1 > 0.2$ or $T_1 < T_x, 1 < x \leq 9$. Otherwise, a lower protective coding scheme may be considered (see Figure 4.6(a)). From all the coding schemes of which the BLER is less than 0.2, the one that has the minimum transfer delay T_x will be chosen as the current coding scheme MCS (see equation Eq.4.6).

$$\begin{aligned} MCS &= MCSx \\ \text{subject to:} & \\ T_x &= \min T_k \qquad \forall\, k|\ BLER_k \leq 0.2 \end{aligned} \tag{4.6}$$

- **The current coding scheme is not MCS1**

  The coding scheme will not change if the following criteria are met:

$$\begin{aligned} BLER_m &\leq 0.2 \\ T_m &= \min T_k \qquad \forall\, k:\ BLER_k \leq 0.2 \end{aligned} \tag{4.7}$$

  Otherwise, another coding scheme will be chosen for the transmission (see Figure 4.6(b)). In that case if the BLER of all the coding schemes is larger than 0.2, only MCS1 can be chosen as the current coding scheme. On the other hand, if there exist the coding schemes of which the BLER is less than 0.2, the one that has the minimum transfer delay T_x will be chosen as the current coding scheme MCS (see equation Eq.4.6).

4.2.4.2 A combined LA and IR scheme

In the combined LA and IR mode,the data block is first transmitted with the MCS (Modulation and Coding Scheme) selected by the LA and after that the IR mechanism takes place. If the transmission of the RLC block is not successful, only the redundant information will be sent by the next retransmission, using a different puncturing scheme of this coding scheme (see subsection 2.3.5.2).

Assume that the radio blocks used for the packet transmission are N_b (see Eq.4.3), the total number of radio blocks N_{total} that has to be used for the transmission can be described as follows:

$$N_{total} = N_b(1{+}BLER(1){+}BLER(1){*}BLER(2){+}BLER(1){*}BLER(2){*}BLER(3)...) \tag{4.8}$$

where $BLER(1)$, $BLER(2)$, $BLER(3)$, $BLER(4)$, ...are the block error rate at LA (no coding combining), IR2 (code combining with P1+P2), IR3 (code combining with P1+P2+P3 or P1+P2+P1), IR4 (code combining with P1+P2+P3+P1 or P1+P2+P1+P2), ...Note that the puncture schemes used for code combining depend on the coding scheme chosen by LA. If all the codewords (different punctured versions of the encoded data block) have been sent, the first codeword is sent, and so on. For instance, for MSC9, the code combining of IR3 is P1+P2+P3, whereas for MCS6, it will be P1+P2+P1. Let M be the total number of retransmissions and i be the i^{th} retransmission, then Eq.4.8 can be written as:

$$N_{total} = N_b(1 + \sum_{i=1}^{M} \prod_{k=1}^{i} BLER(k)) \qquad (M > 0) \tag{4.9}$$

The mean transfer delay T_{delay} will be:

$$T_{delay} = N_{total} \times T = \frac{L}{R_{MCSx}} * (1 + \sum_{i=1}^{M} \prod_{k=1}^{i} BLER(k)) \tag{4.10}$$

Let $BLER\_eff$ represent the effective BLER of the transmission. Compared with Eq. (4.5), Eq. (4.10) can be written as:

$$T_{delay} = \frac{L}{R_{MCSx} * (1 - BLER_{eff})} \tag{4.11}$$

where:

$$BLER\_eff = 1 - \frac{1}{\sum_{n=0}^{N} \prod_{k=0}^{n} BLER(k)} \tag{4.12}$$

Due to the benefits of code combining, the block error rate of the retransmission decreases significantly[47][7]. Assume that $BLER(2)$ is far lower than $BLER(1)$, then Eq. (4.12) simplifies to:

$$BLER\_eff \approx \frac{BLER(1)}{1 + BLER(1)} \tag{4.13}$$

In this case, the decision of the channel coding is to minimize the transfer delay described in Eq.4.11 by considering not to increase excessively the number of retransmissions. We can still assume that the maximum BLER is 20%. Then we have $BLER\_eff_{max} = 0.2$, which means $BLER(1) = 0.25$. We have mentioned that $BLER(1)$ is the block error rate at LA. Therefore, the choice of MCS can be more aggressive after using the combination of LA and IR. The modulation and coding scheme decision based on the BLER and transfer delay is similar to Figure 4.6. However, the BLER changing point increases to 0.25.

4.2.5 Packet Segmentation

The allocation of resources are considered at the block level, i.e. the RLC protocol. So the messages need to be sliced into radio blocks. The process of segmenting messages into radio blocks in EDGE is similar to the procedure in GPRS (see Figure 4.3). In the GPRS simulator, the payload size of a radio block is fixed (181 bits) because only one coding scheme is taken into account. In our simulator, the radio block payload size depends on the coding scheme decided by the LQC algorithm that we are using. For instance, in that case if the current coding scheme is MCS6, the payload size of a radio block is 592.

As we have studied in the last section, the actual number of radio blocks contained in a message depends on the coding scheme, the block error rate and the LQC algorithm.

4.2.6 Packet scheduler

The scheduler assigns turns to each user depending on the queueing weights. It is possible to see the scheduler as a queueing system which consists of queueing, channel allocation and packet scheduling. The implementation details of packet scheduler are not defined in EDGE standard.

In an EDGE network, each new transmission requests a TBF (Temporary Block Flow) to be set up in the RLC/MAC blocks scheduler, and then the TBF is allocated radio resources on one or several physical channels [18]. It is possible to see a TBF as a TBF queue, which contains only the RLC blocks belonging to the same message.

The EDGE standard does not limits the number of simultaneous TBF, but the material implementation of the EDGE network limit the possibilities. Therefore, in our simulator, we assume that a maximum of 4 messages can be simultaneously transferred, which means there are a maximum of 4 TBF queues, each containing the blocks belonging to one message that is being transmitted. We also consider a small waiting buffer which stores arriving messages if there is no resources to set up TBF queues. This buffer only contains one message.

The scheduling algorithm used in this simulator is weighted Iterative Round Robin (weighted IRR). The Weighted Round Robin (WRR) is applied by assigning transmission turns to each class of service. The number of assigned turns is proportional to a weighting factor. The functionality of weighted IRR is as follows: in the first iteration, weighted IRR behaves exactly like WRR. Each user receives a number of logical channels according to its priority level; the low priority users will send only one radio block, while a high priority user will send a weighted number of radio blocks. If there are more channels available after all users sent a number of radio blocks, giving the remaining resources to the users that still have radio blocks waiting in their queue. The algorithm stops when all the available channels were allocated to the users or when all the user's queues are empty.

4.2.7 Providing differentiated services

In our simulator, the differentiation is based on the traffic classes defined in the EDGE standard [29]. Therefore, there are 4 classes of service to be modelled: class

1 represents the users with highest priority (interactive 1) and class 4 the lowest priority (background). According to the EDGE standard, class 4 users do not have the QoS requirements and class 3 users have very low QoS requirements; whereas class 1 and class 2 users have high QoS requirements. For this reason, we just want to propose the mechanisms to guarantee the quality of service of class 1 and class 2 users. We assume that the class importance is weighted as follows: Class 1 services weight 4, class 2 weight 3, class 3 weight 2 and class 4 weight 1. That means class 1 service can use 4 radio blocks each time, while class 4 service can only use one radio block. Weighted IRR is used as scheduling scheme.

In the last subsection, we mentioned TBF queue in our queueing system. A TBF queue is temporarily maintained for the duration of the data transmission and the priority of the queue depends on the message contained in this TBF queue. Based on this characteristic, we proposed the following methods to provide the service to differentiated users:

- **Differentiated queue.** A maximum of 4 messages can be simultaneously transferred, however, the messages belonging to the same service class cannot be transmitted at the same time. The arriving messages will not start their transmission if a message of the same class is being transferred. In this system, the users belonging to class i can only occupy the assigned queue qi. The priority of each queue can be seen as fixed, however, each queue contains only one message.

- **FCFS (First-Come-First-Serve) without priority.** The arriving messages are served following a first-come-first-served policy, regardless of the classes of services. The priority of the TBF queue depends on the message contained in this queue.

- **FCFS with priority.** This method is similar to FCFS without priority. The arrival message is served upon its arrival time without thinking about the class which it belongs to. But the class 1 users have preemptive priority compared with the class 4 users. Assume that there are already 4 messages being transferred when $C1$ arrives at t_n. If $C4$ is being transferred, it will be discarded even if it has not finished its transmission so that $C1$ can be served. The purpose is to reduce the discarding rate of class 1 users. In that case if there are more than one class 4 users being transferred, the one with the longest queue will be replaced.

Before describing the details of these methods, let us suppose the following example. We assume that the allocation of resources will be updated every duration T_r (consists of 16 radio blocks) and that the total available logical channels are 16. Let us also assume that each message is transmitted in 24 radio blocks. Weighted IRR is used as scheduling scheme. Let Cm_{tn} be a user of class m arriving at time tm, m=0,1,2,...,10, and $t0 < t1 < t2... < t11$ (see table 4.1).

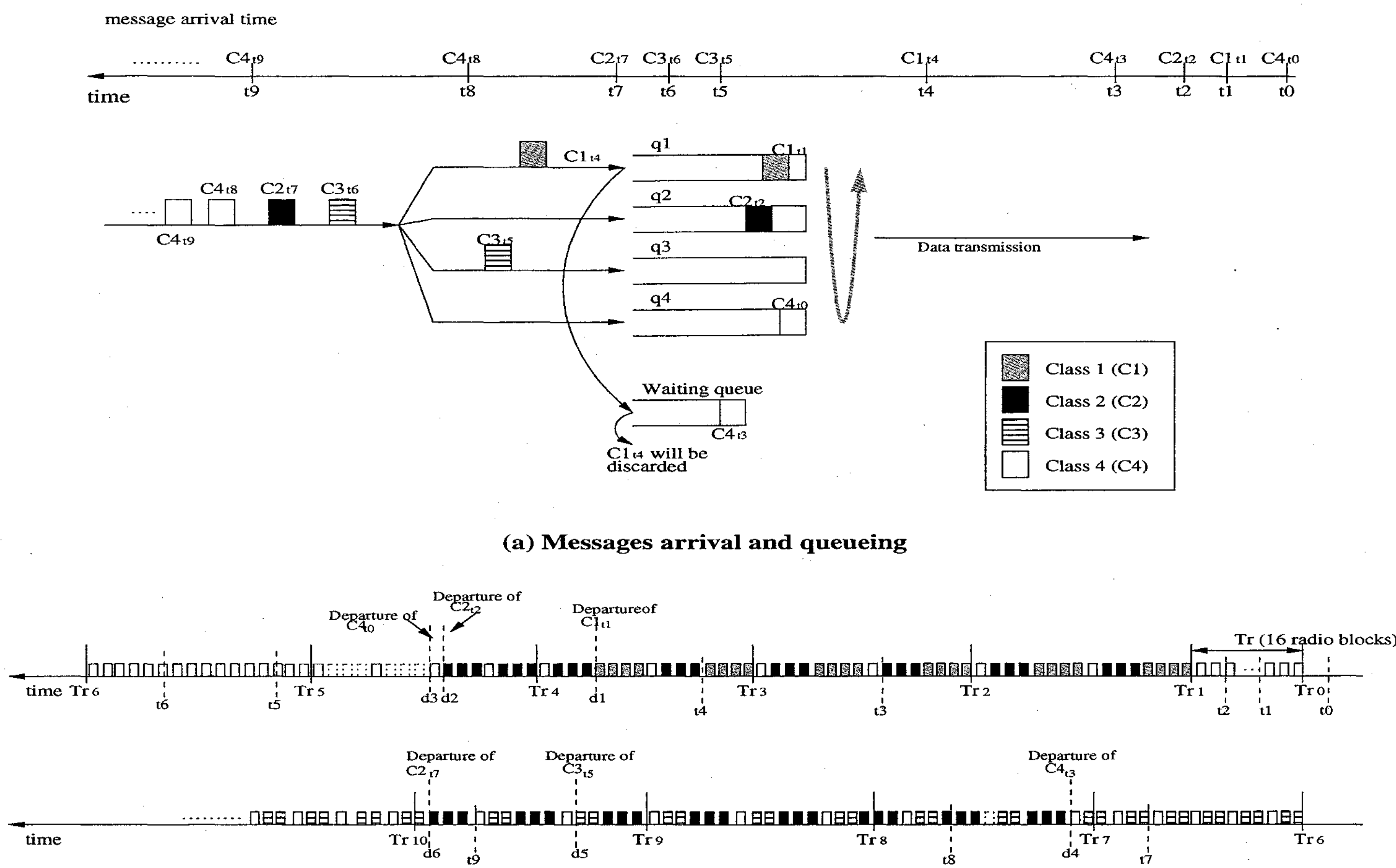
Table 4.1: Packet arrival time example

Arrival time	$t0$	$t1$	$t2$	$t3$	$t4$	$t5$	$t6$	$t7$	$t8$	$t9$
Class of service	$C4_{t0}$	$C1_{t1}$	$C2_{t2}$	$C4_{t3}$	$C1_{t4}$	$C3_{t5}$	$C3_{t6}$	$C2_{t7}$	$C4_{t8}$	$C4_{t9}$

Now we can talk about the details of these method.

- **Differentiated queue**

  Considering the example we have mentioned, let qi represent the TBF queue set up by a class i user, the blocks of $C4_{t0}$ will be contained in $q4$, the blocks of $C1_{t1}$ in $q1$, the blocks $C2_{t2}$ in $q2$. Then when $C4_{t3}$ arrives, it will be put in the waiting queue instead of being served even if only three messages are being transmitted (see Figure 4.7(a)).



(a) Messages arrival and queueing

(b) Transmission process

Figure 4.7: Differentiated queue

The transmission process is portrayed in Figure 4.7(b), where d_i denotes the i^{th} departure from the system and T_{rj} represents the j^{th} beginning of the update of the resource allocation. Table 4.2 shows the interval of departure time for each user.

Table 4.2: Packet departure time for differentiated queue

User	Departure time
$C1_{t1}$	d_1 ($t4 < d_1 < t5$)
$C2_{t2}$	d_2 ($t4 < d_2 < t5$)
$C4_{t0}$	d_3 ($t4 < d_3 < t5$)
$C4_{t3}$	d_4 ($t7 < d_4 < t8$)
$C3_{t5}$	d_5 ($t8 < d_5 < t9$)
$C2_{t7}$	d_6 ($d_6 > t9$)

The transmission process is described as follows:

- **Time period between T_{r0} and T_{r1}.**
  $C4_{t0}$ arrives at $t0$ and is sent to $q4$. At time T_{r0}, $C4_{t0}$ is assigned transmission turns. Note that all the 16 logical channels are assigned to $C4_{t0}$ since the other three queues are empty. At $t1$ and $t2$ respectively, $C1_{t1}$ and $C2_{t2}$ arrive and are sent to $q1$ and $q2$ respectively. However, they cannot start their transmission until the resource allocation is updated.

- **Time period between T_{r1} and T_{r2}.**
  At T_{r1}, the pointer goes to $q1$ and the resources are re-allocated. In the first iteration, according to their weights, $C1_{t1}$, $C2_{t2}$ and $C4_{t0}$ are assigned 4, 3, and 1 logical channels respectively. The allocation continues until all the channels are assigned to the users. The transmission sequence is as follows (the number in the parentheses indicates the transmission

turns assigned to this user): $C1_{t1}(4)$, $C2_{t2}(3)$, $C4_{t0}(1)$, $C1_{t1}(4)$, $C2_{t2}(3)$, $C4_{t0}(1)$.

- **Time period between T_{r2} and T_{r4}.**

  From T_{r2} to T_{r4}, the allocation is the same as the last period. Note that when $C4_{t3}$ arrives at $t3$, it is sent to the waiting queue since $C4_{t0}$ is being transferred and it cannot be served. Then, when $C1_{t4}$ arrives at $t4$, $C1_{t1}$ has not finish transmitting and the waiting queue is full. Hence, $C1_{t4}$ has to be discarded even if there is the place for the transmission. $C1_{t1}$ finishes its transmission at d_1. The transmission sequence in this period is: $C1_{t1}(4)$, $C2_{t2}(3)$, $C4_{t0}(1)$, $C1_{t1}(4)$, $C2_{t2}(3)$, $C4_{t0}(1)$, ...

- **Time period between T_{r4} and T_{r5}.**

  At T_{r4}, only $C2_{t2}$ and $C4_{t0}$ have the blocks waiting in their queues. Thus, the transmission turns are proportionally assigned to $C2_{t2}$ and $C4_{t0}$. After finishing the transmission, $C2_{t2}$ and $C4_{t0}$ depart at d_2 and d_3 respectively. And then $C4_{t3}$ will be sent to $q4$ and use the transmission turns assigned to $q4$. At this time, $q2$ is empty, but there are still the transmission turns assigned to it. We synchronize radio resources by sending empty blocks to fill the gaps. The transmission sequence in this period is: $C2_{t2}(3)$, $C4_{t0}(1)$, $C2_{t2}(3)$, $C4_{t0}(1)$,*empty*(3), $C4_{t3}(1)$,*empty*(3), $C4_{t3}(1)$.

- **Time period between T_{r5} and T_{r6}.**

  All the 16 logical channels are assigned to $C4_{t3}$. During this period, $C3_{t5}$ arrives at $t5$ and is sent to $q3$, whereas $C3_{t6}$ arrives at $t6$ and has to be sent to the waiting queue.

- **Time period between T_{r6} and T_{r7}.**

  At T_{r6}, The pointer goes to $q3$ and the transmission turns are allocated to $q3$ and $q4$ respectively. $C2_{t7}$ arrives at $t7$ and is sent to $q2$ to wait for the transmission turn. The transmission sequence in this period is: $C3_{t5}(2)$,

$C4_{t3}(1)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t5}(1)$.

- **Time period between T_{r7} and T_{r8}.**

  At T_{r7}, The pointer goes to $q3$ and the transmission turns are allocated to $q3$, $q4$ and $q2$ respectively. $C4_{t3}$ departs at d_4. Therefore, when $C4_{t8}$ arrives at $t8$, it will be sent to $q4$ and use the transmission turns assigned to $q4$. The transmission sequence in this period is: $C3_{t5}(1)$, $C4_{t3}(1)$, $C2_{t7}(3)$, $C3_{t5}(2)$, $empty(1)$, $C2_{t7}(3)$, $C3_{t5}(2)$, $C4_{t8}(1)$, $C2_{t7}(2)$.

- **Time period between T_{r8} and T_{r9}.**

  At T_{r8}, The pointer goes to $q2$ and the transmission turns are allocated to $q2$, $q3$ and $q4$ respectively. The transmission sequence in this period is: $C2_{t7}(1)$, $C3_{t5}(2)$, $C4_{t8}(1)$, $C2_{t7}(3)$, $C3_{t5}(2)$, $C4_{t8}(1)$, $C2_{t7}(3)$, $C3_{t5}(2)$, $C4_{t8}(1)$.

- **Time period between T_{r9} and T_{r10}.**

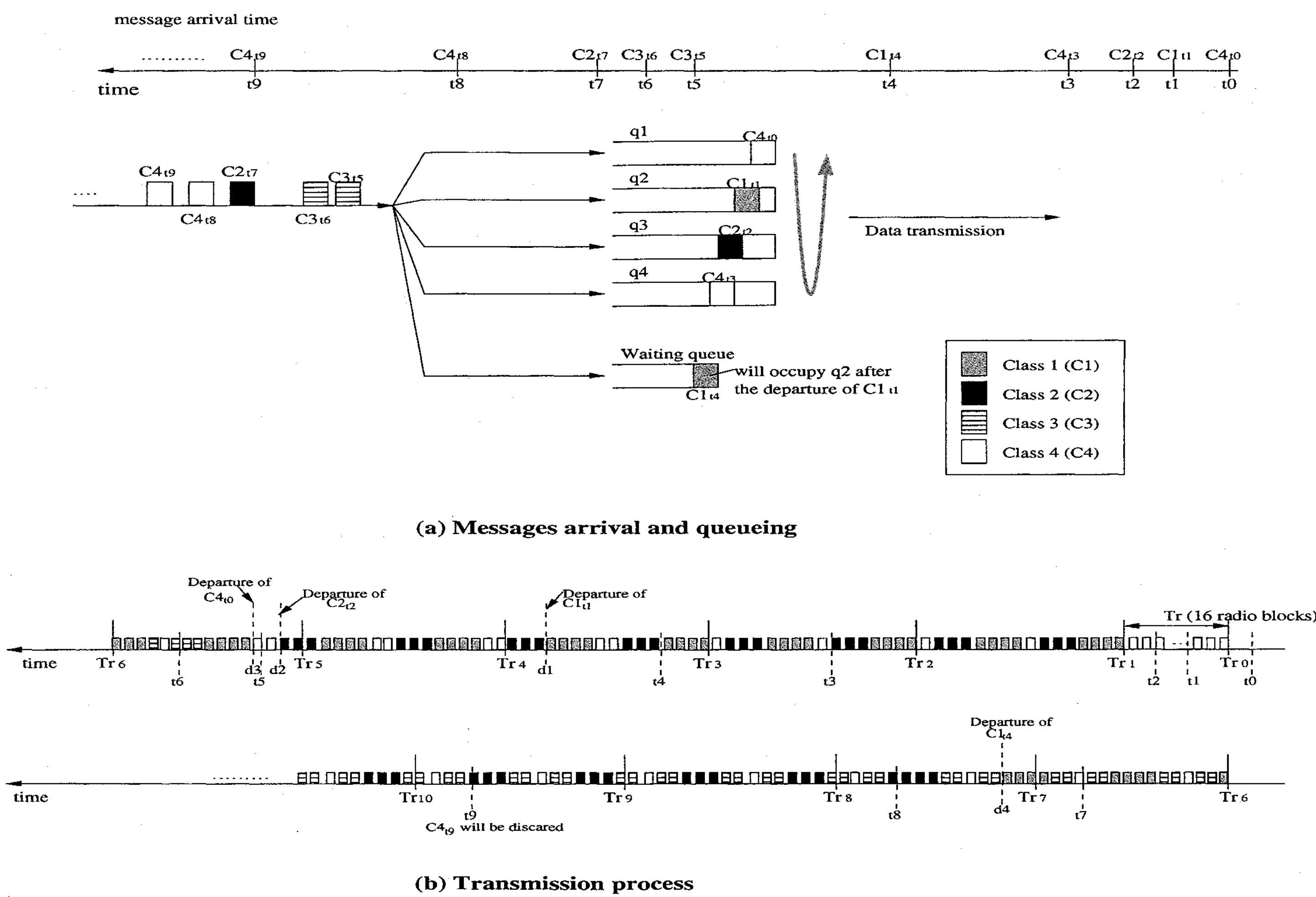
  The transmission turns are allocated to $q2$, $q3$ and $q4$ respectively. After the departure of $C3_{t5}$, $C3_{t6}$ will be sent to $q3$ and use the transmission turns assigned to $q3$. When $C4_{t9}$ arrives at $t9$, it will be put into the waiting queue. The transmission sequence in this period is: $C2_{t7}(3)$, $C3_{t5}(2)$, $C4_{t8}(1)$, $C2_{t7}(3)$, $C3_{t6}(2)$, $C4_{t8}(1)$, $C2_{t7}(3)$, $C3_{t5}(1)$.

The process continues until all the users finish their transmission.

- **FCFS without priority**

  Repeat the example given above, let qi represents the TBF queue set up by an arrival message, the blocks of $C4_{t0}$ will be put in $q1$, the blocks of $C1_{t1}$ in $q2$, the blocks of $C2_{t2}$ in $q3$ and the blocks of $C4_{t3}$ in $q4$ (see Figure 4.8(a)).

  The transmission process is portrayed in Figure 4.8(b). And the interval of departure time for each user is shown in table 4.3.



(a) Messages arrival and queueing

(b) Transmission process

Figure 4.8: FCFS without priority

Table 4.3: Packet departure time for FCFS without priority

User	Departure time
$C1_{t1}$	d_1 ($t4 < d_1 < t5$)
$C2_{t2}$	d_2 ($t4 < d_2 < t5$)
$C4_{t0}$	d_3 ($t5 < d_3 < t6$)
$C1_{t4}$	d_4 ($t7 < d_4 < t8$)

The transmission process is described as follows:

- **Time period between T_{r0} and T_{r1}.**

  The process is identical to the one in the previous method. $C4_{t0}$ arrives at $t0$ and the blocks are sent to $q1$. At time T_{r0}, $C4_{t0}$ is assigned transmission turns and all the 16 logical channels are assigned to $C4_{t0}$. At $t1$ and $t2$ respectively, $C1_{t1}$ and $C2_{t2}$ arrive and their blocks are sent to $q2$ and $q3$ respectively. However, they cannot start their transmission until the resource allocation is updated.

- **Time period between T_{r1} and T_{r2}.**

  At T_{r1}, the pointer goes to $q2$ and the transmission turns are assigned to $q2$, $q3$ and $q1$ respectively. The transmission sequence is as follows (the number in the parentheses indicates the transmission turns assigned to this user): $C1_{t1}(4)$, $C2_{t2}(3)$, $C4_{t0}(1)$, $C1_{t1}(4)$, $C2_{t2}(3)$, $C4_{t0}(1)$.

- **Time period between T_{r2} and T_{r3}.**

  The pointer goes to $q2$ and the allocation is the same as last period. When $C4_{t3}$ arrives at $t3$, its blocks are sent to $q4$ to wait for the transmission turn. The transmission sequence in this period is: $C1_{t1}(4)$, $C2_{t2}(3)$, $C4_{t0}(1)$, $C1_{t1}(4)$, $C2_{t2}(3)$, $C4_{t0}(1)$.

- **Time period between T_{r3} and T_{r4}.**

  The pointer goes to $q2$ and the transmission turns are assigned to $q1$, $q2$, $q3$ and $q4$ respectively. When $C1_{t4}$ arrives at $t4$, it is put into the waiting queue. After the departure of $C1_{t1}$ (at d_1), the blocks of $C1_{t4}$ will be put into $q2$. The transmission sequence in this period is: $C1_{t1}(4)$, $C2_{t2}(3)$, $C4_{t3}(1)$, $C4_{t0}(1)$, $C1_{t1}(4)$, $C2_{t2}(3)$.

- **Time period between T_{r4} and T_{r5}.**

  At T_{r4}, the pointer goes to $q4$ and the transmission turns are assigned to the users according to their weights. At this time, $C4_{t0}$ is in $q1$, $C1_{t4}$ in $q2$, $C2_{t2}$ in $q3$ and $C4_{t3}$ in $q4$. The transmission sequence in this period is: $C4_{t3}(1)$, $C4_{t0}(1)$, $C1_{t4}(4)$, $C2_{t2}(3)$, $C4_{t3}(1)$, $C4_{t0}(1)$, $C1_{t4}(4)$, $C2_{t2}(1)$.

- **Time period between T_{r5} and T_{r6}.**

  The transmission turns assigned to each transfer queue are the same as last period. $C2_{t2}$ departs at d_2 so that when $C3_{t5}$ arrives, its blocks can be put into $q3$ and use the transmission turns assigned to $q3$. $C4_{t0}$ departs at d_3. Therefore, when $C3_{t6}$ arrives, it is sent to $q1$ and uses the transmission turns assigned to $q1$. The transmission sequence in this period is: $C2_{t2}(2)$, $C4_{t3}(1)$, $C4_{t0}(1)$, $C1_{t4}(4)$, $C3_{t5}(3)$, $C4_{t3}(1)$, $C3_{t6}(1)$, $C1_{t4}(3)$.

- **Time period between T_{r6} and T_{r7}.**

  At T_{r6}, The pointer goes to $q2$ and the transmission turns are allocated to each transfer queue according to their weights. $C2_{t7}$ arrives at $t7$ and is put into the waiting queue. The transmission sequence in this period is: $C1_{t4}(1)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t6}(2)$, $C1_{t4}(4)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t6}(2)$, $C1_{t4}(1)$.

- **Time period between T_{r7} and T_{r8}.**

  The pointer goes to $q2$ and the transmission turns assigned to each transfer queue are the same as last period. After the departure of $C1_{t4}$ at d_4, $C2_{t7}$

will occupy $q2$ and use the transmission turns assigned to $q2$. $C4_{t8}$ arrives at $t8$ and is put into the waiting queue. The transmission sequence in this period is: $C1_{t4}(3)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t6}(2)$, $C2_{t7}(4)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t6}(1)$.

- **Time period between T_{r8} and T_{r9}.**

  At T_{r8}, The pointer goes to $q1$ and the transmission turns are re-allocated. The transmission sequence in this period is: $C3_{t6}(1)$, $C2_{t7}(3)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t6}(2)$, $C2_{t7}(3)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t6}(1)$.

- **Time period between T_{r9} and T_{r10}.**

  The pointer goes to $q2$ and the transmission turns assigned to each transfer queue are the same as last period. When $C4_{t9}$ arrives at $t9$, it will be dropped since the resources are unavailable. The transmission sequence in this period is: $C3_{t6}(1)$, $C2_{t7}(3)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t6}(2)$, $C2_{t7}(3)$, $C3_{t5}(2)$, $C4_{t3}(1)$, $C3_{t6}(1)$.

The process continues until all the users finish their transmission.

- **FCFS with priority**

  Using again the given example, the blocks of $C4_{t0}$ will be put in $q1$, the blocks of $C1_{t1}$ in $q2$, the blocks of $C2_{t2}$ in q3 and the blocks of $C4_{t3}$ in $q4$ (see Figure 4.9(a)).

  The transmission process is portrayed in Figure 4.9(b). The interval of departure time for each user is shown in table 4.3.

  The transmission process is described as follows:

  - **Time period between T_{r0} and T_{r3}.**

    during these periods, the transmission process is identical to the one in the previous method (FCFS without priority). $C4_{t0}$, $C1_{t1}$, $C2_{t2}$ and $C4_{t3}$

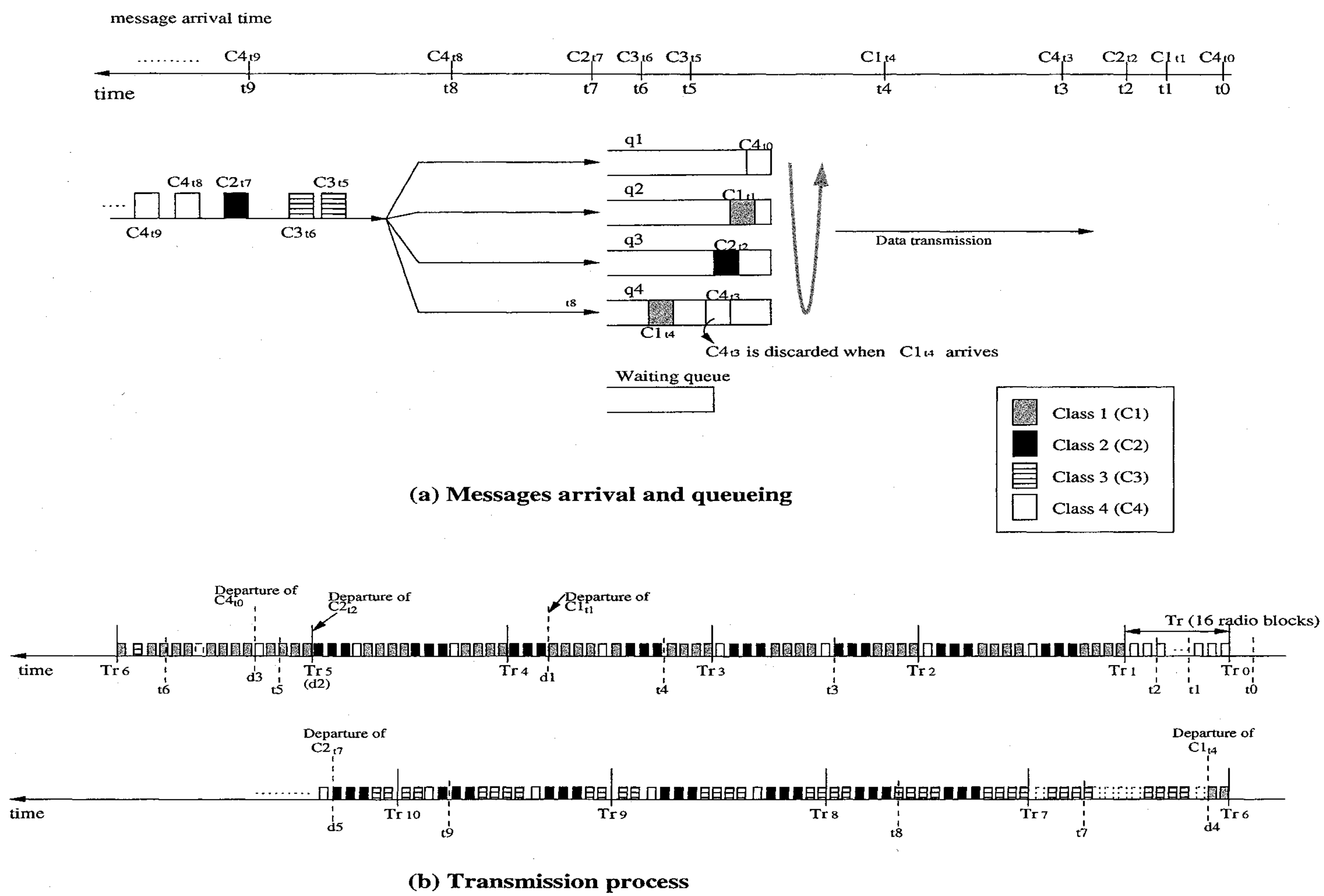


Figure 4.9: FCFS with priority

Table 4.4: Packet departure time FCFS with priority

User	Departure time
$C1_{t1}$	d_1 ($t4 < d_1 < t5$)
$C2_{t2}$	d_2 ($t4 < d_2 < t5$)
$C4_{t0}$	d_3 ($t4 < d_3 < t5$)
$C1_{t4}$	d_4 ($t6 < d_4 < t7$)
$C2_{t7}$	d_5 ($d_5 > t9$)

occupy $q1$, $q2$, $q3$ and $q4$ respectively. The transmission turns are assigned to different queues according to the weights of the users.

- **Time period between T_{r3} and T_{r4}.**

  At T_{r3}, the pointer goes to $q2$ and the transmission turns are assigned to $q1$, $q2$, $q3$ and $q4$ respectively. When $C1_{t4}$ arrives at $t4$, there are already 4 messages being transferred. However, $C4_{t0}$ and $C4_{t3}$ are being transferred and $q4$ is longer than $q3$. Therefore, $C4_{t3}$ will be discarded so that $C1_{t4}$ use this place and use the transmission turns assigned to $q4$. $C1_{t1}$ departs at d_1. The transmission sequence in this period is: $C1_{t1}(4)$, $C2_{t2}(3)$, $C1_{t4}(1)$, $C4_{t0}(1)$, $C1_{t1}(4)$, $C2_{t2}(3)$.

- **Time period between T_{r4} and T_{r5}.**

  At T_{r4}, the pointer goes to $q4$ and the transmission turns are assigned to the users according to their weights. At this time, $C4_{t0}$ is in $q1$, $C2_{t2}$ in $q3$ and $C1_{t4}$ in $q4$. $C2_{t2}$ will finish its transmission and depart at d_2. The transmission sequence in this period is: $C1_{t4}(4)$, $C4_{t0}(1)$, $C2_{t2}(3)$, $C1_{t4}(4)$, $C4_{t0}(1)$, $C2_{t2}(3)$.

- **Time period between T_{r5} and T_{r6}.**

  At T_{r5}, the pointer goes to $q4$ and the transmission turns are assigned to $q4$ and $q1$. when $C3_{t5}$ arrives at $t5$, it will occupy $q2$ to wait for the transmission turn. After finishing the transmission, $C4_{t0}$ departs at d_3. Thus, when $C3_{t6}$ arrives at $t5$, it will be sent to $q1$ and can use the transmission turns assigned to $q1$. The transmission sequence in this period is: $C1_{t4}(4)$, $C4_{t0}(1)$, $C1_{t4}(4)$, $empty(1)$, $C1_{t4}(4)$, $C3_{t6}(1)$, $C1_{t4}(1)$.

- **Time period between T_{r6} and T_{r7}.**

  At T_{r6}, The pointer goes to $q4$ and the transmission turns are allocated to $q4$, $q1$ and $q2$ respectively. $C1_{t4}$ finishes the transmission and departs at d_4. Then the empty blocks will be sent to synchronize the radio resources. $C2_{t7}$ arrives at $t7$ and is sent to $q3$ to wait for the transmission turn. The transmission sequence in this period is: $C1_{t4}(2)$, $empty(1)$, $C3_{t6}(2)$, $C3_{t5}(2)$, $empty(4)$, $C3_{t6}(2)$, $C3_{t5}(2)$, $empty(1)$.

- **Time period between T_{r7} and T_{r8}.**

  The pointer goes to $q1$ and the transmission turns are assigned to $q1$, $q2$ and $q3$ respectively. $C4_{t8}$ arrives at $t8$ and its blocks are put into $q4$. The transmission sequence in this period is: $C3_{t6}(2)$, $C3_{t5}(2)$, $C2_{t7}(3)$, $C3_{t6}(2)$, $C3_{t5}(2)$, $C2_{t7}(3)$, $C3_{t6}(2)$.

- **Time period between T_{r8} and T_{r9}.**

  At T_{r8}, The pointer goes to $q2$ and the transmission turns are re-allocated. The transmission sequence in this period is: $C3_{t5}(2)$, $C2_{t7}(3)$, $C4_{t8}(1)$, $C3_{t6}(2)$, $C3_{t5}(2)$, $C2_{t7}(3)$, $C4_{t8}(1)$, $C3_{t6}(2)$,

- **Time period between T_{r9} and T_{r10}.**

  When $C4_{t9}$ arrives at $t9$, it will be put into the waiting queue. The transmission sequence in this period is: $C3_{t5}(2)$, $C2_{t7}(3)$, $C4_{t8}(1)$, $C3_{t6}(2)$, $C3_{t5}(2)$, $C2_{t7}(3)$, $C4_{t8}(1)$, $C3_{t6}(2)$,

The process continues until all the users finish their transmission.

4.3 Implementation

The simulation model is implemented using the C++ language and the library of simulation called Sim written by D.Boliers and A.Eliens from Amsterdan faculty of science. The Sim library is a C++ library used for discrete events simulation. the components of the model are events, which are activated at certain points in time and in this way affect the overall state of the system. The moments that events are activated are randomly produced according to a certain distribution. The library contains a simulation class for running and ending the simulation, a primitive class for managing the events, a queue class with priority, a histogram class for gathering results and a generator class for producing the stochastic distributions.

When a simulation is created, it runs after invoking the *simulation::run* method in session *simulation::main* function. The simulation then runs until there are no events left or for a specified simulation time. The general scheme of the process of a simulation is depicted in Figure 4.10.

For example, let us see how a M/G/C/K queue is modelled by this library. The first event is created at the beginning of the simulation, then the next event is created during the activation of the previous event. A new user is given an activation moment from an exponential distribution. A user activation happens when it enters the queue, so the arrival process is Poisson. If there are already K events in the queue, new events are activated and terminated right away. Among the C servers, the free ones look for users in the queue and apply a service time. The distribution of the service time depends on the users. For instance, geometric, normal, etc. At the end we have the simulation of a M/G/C/K queue. For more details on the sim library, see the official sim manual in http://www.cs.vu.nl/ eliens/sim/.

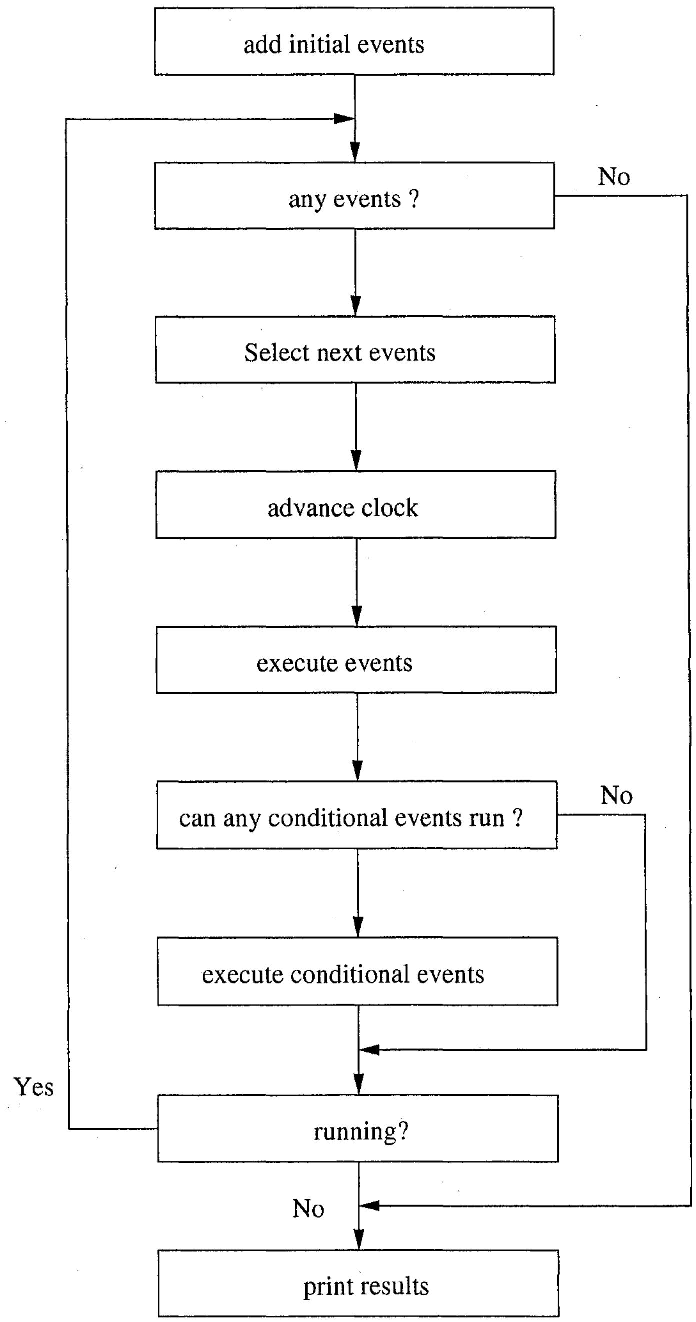


Figure 4.10: General process of the execution of a simulation employed by sim

Chapter 5

Simulation results

We have concentrated our study in the downlink performance evaluation. The reason is that comparing with the uplink transmission, the downlink communication carries much greater volume of information from the servers outside of the wireless networks, e.g. WWW servers and Email servers.

5.1 QoS measures

We have chosen to study the same user QoS measures proposed in [32] since these measures can generally represent the quality of data service in a network. They were also the measures favored by the network operator on the work presented in [32]. They are:

- **The delivery time**. This measure reflects the total time that a user have to wait to transfer his message. We will provide two kinds of measures: the mean delivery time and the maximum delivery time for 95 % of the users. Here **the maximum delivery time for 95 % of the users** means that 95 % of the users will experience less delivery time than this value.

- **The delivery time for the first block of data**. This measure provides the user with an assessment on the response time of the system.
- **The blocking rate**. This measure gives the user a measure of the availability of the system.

The quality of service is measured at the message level. Thus, we evaluate the delivery time of a message as the arrival time of the last radio block of the message minus the departure time of the first radio block.

5.2 approximations

Before running the simulation, we make a series of assumptions and modelling approximations.

5.2.1 Connection delay

Before being able to start the data transmission, there is a procedure such as downlink assignment, packet control acknowledgement, etc[18]. For the moment, the delay introduced by the connection protocol is not considered in the simulation. But it is possible to add this parameter in our simulator.

5.2.2 Channel quality

For the same user, we assume that the channel quality of each physical channel will be the same. And the channel quality will not change during the transmission of a message.

5.2.3 Transmission errors

The transmission errors are represented by the block error rate (BLER), whereas the BLER will be implemented by an increase in the effective radio block numbers that should be used to finish a packet transmission. Let L be the message length, T the mean transmission time per radio block ($T = 20ms$) ,R_{MCSx} the data rate of a given coding scheme (see table 2.1) and $BLER_{MCSx}$ the block error rate of a given coding scheme at a given C/I without code combining. In that case if we use LA as LQC algorithm, the total radio block numbers N_{total} that should be used to transmit a message will be obtained by Eq. (5.1) (see section 4.2.4.1). On the other hand, if we use LA+IR as LQC algorithm, the actual numbers can be obtained by Eq. (5.2) (see section 4.2.4.2).

$$N_{total} = \frac{L}{T * R_{MCSx} * (1 - BLER_{MCSx})} \tag{5.1}$$

$$N_{total} = \frac{L}{T * R_{MCSx} * (1 - BLER\_eff)} \tag{5.2}$$

where:

$$BLER\_eff \approx \frac{BLER_{MCSx}}{1 + BLERMCSx} \tag{5.3}$$

Let us consider the following example:

- The message length is 50kB
- The current coding scheme is MCS6, which means $R_{MCSx} = 29.6$kbits/s
- The block error rate is 0.15

Then if LA is used, the actual radio block numbers will be:$N_{total} = \frac{50*1024*8}{592*(1-0.15)} = 814$. On the other hand, if a combined LA and IR is used, $BLER\_eff$ will be 0.13 and N_{total} will be 796.

5.2.4 Overhead and compression

As we have seen, the segmentation from message data to radio block will pass through several protocol layers (see Figure 5.1). Each layer will change the total size of the transmitted data.

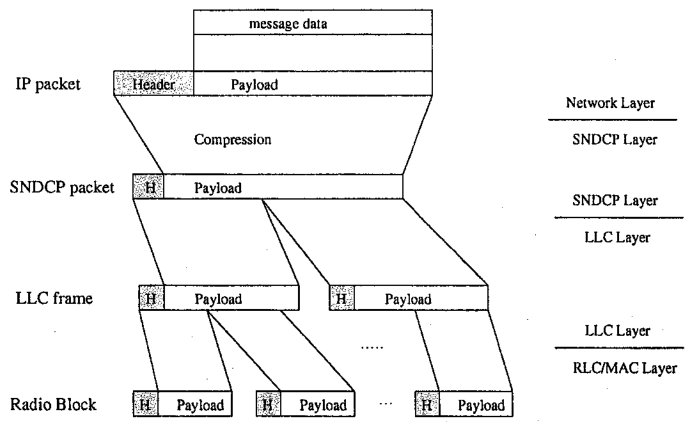


Figure 5.1: Protocol layers used for segmentation

A TCP/IP header will be added before the data enters the SNDCP layer. Then the SNDCP layer receives the IP packet and makes the compression. The comression includes the header compression and data compression, but the data compression is optional. The GSM standard [30] defined that the header compression follows the protocol RFC 1144. Accoding to RFC 1144[33], The raw data expands by 19% when encapsulated in TCP/IP. and by 2% when encapsulated in header compressed TCP/IP. Each LLC frame consists of the header, the trailer, and the information field. The size of the header and tailer are variable but typically consist of 5 to 7 octets for each LLC frame[20]. The LLC frame size is variable from 20 to 1500 octets. In RLC layer, the LLC frame will be segmented into radio blocks. Then information bits contained in each radio block depends on the coding scheme (see table 2.1). The overhead and compression effects introduced by SNDCP, LLC and the RLC/MAC layer are taken into account by the increase of the mean transmission time of each radio block. For instance, let us consider the following example:

- The size of the message that needs to be transmitted is L kB.
- Only the header compression is considered.
- The size of the LLC frame is 1024 octets, i.e. the number of LLC frames contained in the message is L.
- The LLC header and tailer consist of 7 octets.

The message data encapsulated in header compressed TCP/IP will cause an increase of 2% of the data size and each LLC frame add 7 octets. Therefore, the total data size is $L * 1024 * 8 * 1.02 + 7 * 8 * L$, and the increase rate given by the standard is 1.027. The transmission time of each radio block is 0.02s [18](see subsection 2.3.2). When considering the effect of header and compression, the equivalent transmission time of a radio block is $0.02 * 1.027 = 0.0205s$.

5.3 Simulation Parameters

The following parameters will be used in our simulation to obtain the results:

- The messages inter-arrival time and the voice inter-arrival time follow a Poisson distribution. The mean inter-arrival time for messages is from 10 seconds to 25 seconds to represent different system load, whereas for voice calls is 15 seconds (0.06 calls/s). In fact, the distribution of message inter-arrival time can be a different distribution when using different input parameters, e.g. Weibull, hyperexponential, etc, which makes it possible to simulate different traffic models.
- Voice call duration distribution is assumed to be exponential with mean 60 seconds.

- Mean transmission time of a radio block is 0.0205s.
- Message lengths are distributed as follows: 15% present 1KB, 25% 2KB, 15% 5KB, 13% 10KB, 12% 15KB, 5% 50KB, 4 % 100KB, 4% 500KB, 2% 1000KB.
- There are two logical channels for each physical channel. Logical channel 1 is assigned to blocks B0, B2, B4, B6, B8, B10 whereas logical channel 2 is assigned to B1, B3, B5, B7, B9, B11 (see Figure 5.2).

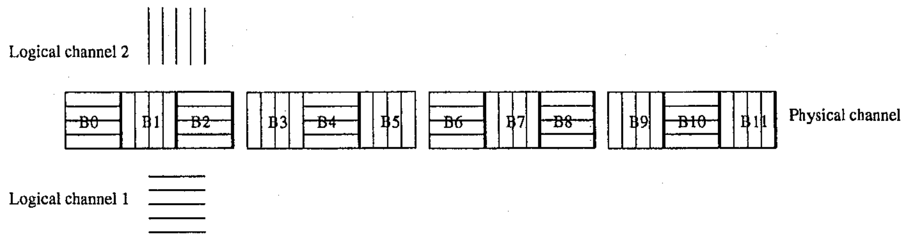


Figure 5.2: Logical channels used in the simulation

- The total physical channels are eight, one channel is for signalling, three channels are reserved for voice, a maximum of four channels can be used for data transmission.
- The channel type is TU3 (typical urban channel with mobile station speed of 3 km/h) without frequency hopping. At a given C/I, the BLER is different with different channel type (e.g. TU3,TU50)[23]. And the block error rate will be calculated according to the data mentioned in GSM 5.05[23] and 3GPP Tdoc[1].
- The carrier-to-interference ratio (C/I) follows a Log-Normal distribution with the mean of 16dB and standard deviation of 7dB[49].

5.4 Confidence interval

A confidence interval is an interval used to estimate the likely size of a population parameter. It provides an indication of the error associated with the sample mean. The confidence interval represents a range of values around the sample mean that include the true mean. In other words, a confidence interval can be represented as a range Δ, such that $Pr\{\bar{x} - \Delta < x < \bar{x} + \Delta\} = 1 - a$, where $\bar{x}$ = the sample mean, Δ = standard error, $1 - a$ = the level of the confidence interval.

In order to obtain the confidence interval for the sample mean we have to first estimate the standard deviation. If the samples $x_1, x_2, ..., x_n$ are independent of each other, then the sample mean is:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{5.4}$$

And the standard deviation δ is:

$$\delta = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{5.5}$$

or, using the short-cut mode:

$$\delta = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} x_i^2 - \frac{(\sum x_i)^2}{n}} \tag{5.6}$$

In our simulation, the confidence interval of mean delivery time is considered and we use the confidence interval at 95% confidence. That is, If samples of the same size are drawn repeatedly from a population, and a confidence interval is calculated from each sample, then 95% of these intervals should contain the population mean. For large samples (i.e. n>30), the confidence interval at 95% confidence is:

$$(\bar{x} - 1.96\frac{\delta}{\sqrt{n}}, \bar{x} + 1.96\frac{\delta}{\sqrt{n}}) \tag{5.7}$$

For small samples, we can construct our confidence interval using the student distribution. For example, let $n = 9$, then the confidence interval at 95% confidence is:

$$(\bar{x} - 2.036\frac{\delta}{\sqrt{n}}, \bar{x} + 2.036\frac{\delta}{\sqrt{n}}) \tag{5.8}$$

In order to get a more precise assessment of the long-term behavior of a system, each simulation case is going to be rerun nine times. Each time, the random generator seeds will be changed in order to generate different events for the simulation. Then, the results over all the nine cases will be considered to calculate the mean delivery time and confidence interval as the long-term results of our simulation. On the other hand, the results on the maximum delivery time for 95% of the users, the blocking rate and the first block delivery time can be obtained by the mean results over the nine cases.

5.5 Simulation results

First of all, we will present the results to see the effects of link quality control. Then, we will present the results based on the different policies of network technologies in order to find the factors that influence the QoS offered to the users. These technologies consist of the data channel reservation policies, the different LQC algorithms and the admission control. Finally, the results based on the different QoS policies will be presented. These policies include the QoS policy based on the message size and the QoS policy based on the differentiated users. We try to find if the use of the QoS policy can improve the quality of service offered to the users. Therefore, the following six types of results will be presented:

- **Results for effects of Link Quality Control**
- **Results for different data channel reservation policies**

- **Results for different LQC Algorithms**
- **Results on the effects of admission control**
- **Results for QoS policy based on message size**
- **Results for a type of QoS policy based on differentiated users**

For the first four types of results, the classes of service are not taken into account, whereas for the last two, the users are grouped using the different QoS policies. The following sections will describe the details of these six types of results.

5.6 Results for effects of Link Quality Control

We have emphasize many times the importance of link quality control in an EDGE network. Without LQC, only the most robust coding scheme (e.g. MCS1) can be used to satisfy the wide area coverage (see Figure 2.1). In this section, we will compare the QoS offered to user in an EDGE network with and without LQC. In order to show the enhancement of EDGE, we will also compare the results obtained in GPRS simulator proposed in [32]. The simulation parameters used in the GPRS simulator are the same as those used in the EDGE simulator. However, the GPRS simulator does not consider the channel condition. According to the distribution of C/I that we used in the EDGE simulator, the average BLER of the most robust GPRS coding scheme (CS1) is 0. Therefore, this value is used to get the results of the GPRS simulator.

The other parameters are as follows:

Admission control is not taken into account. All the users will be accepted by the system if the radio resources are available. Two physical channels are reserved for data transmission. For the results without LQC, MCS1 is used as coding scheme. On the other hand, for the results with LQC, link Adaptation is used as LQC algorithm.

The classes of service are not taken into account.

The simulation results are portrayed in appendix A. In Figures A.1 to A.10, the mean delivery time of each message size in different cases are compared, whereas in Figures A.11 to A.20, the maximum delivery time for 95% of the users are portrayed. The X_axis represent the system load and the Y_axis shows the delivery time. Tables 5.1 and 5.2 also list the delivery time of each message size in different cases. In our simulation, the range of data traffic load is from 0.04 messages/s to 0.1 messages/s. However, only the results of two traffic loads are written in the tables for the QoS comparison and the best values are emphasized in bold.

As we can see, without LQC, the performance for EDGE network is even worse than GPRS network because the transfer rate of MCS1 is lower than CS1 (8.8 kbits/s vs 9.05 kbits/s). However, when LQC is used, both the mean delivery time and the maximum delivery time for 95% of the users in an EDGE network are much lower than the other two cases. Figures A.21 and A.22 present the message discard rate and delivery time for the first block in these three cases. We can also find the numerical results in table 5.4 and table5.3 and the best results are portrayed in bold in the tables.

As expected, after using the LQC algorithm, the message discarding rate and the first block delivery time decrease obviously. In a word, using of LQC can significantly improve the QoS offered to users.

5.7 Results for different data channel reservation policies

In this section, we present the comparison for three different types of reservation policies:

Table 5.1: QoS comparison: mean delivery time with or without LQC (s)

System Load	Case	Message size (kB)									
		1	2	5	10	15	30	50	100	500	1000
0.04 msg/s	GPRS	1.74±0.67	2.33±0.54	4.22±0.80	6.73±0.75	10.02±0.94	18.65±1.70	28.91±1.85	56.43±3.46	274.22±12.55	539.73±29.57
	EDGE (no LQC)	2.68±0.33	3.44±0.26	5.52±0.34	9.29±0.38	12.60±0.37	23.47±0.79	37.76±1.0	71.86±1.62	347.52±6.22	693.95±16.83
	EDGE (LQC)	**0.52±0.04**	**0.82±0.07**	**1.60±0.07**	**2.94±0.09**	**4.28±0.12**	**8.16±0.32**	**13.37±0.5**	**26.74±1.08**	**125.78±4.53**	**251.46±13.05**
0.1 msg/s	GPRS	6.92±1.23	7.44±0.87	10.53±1.19	15.07±1.29	19.56±1.37	33.10±2.29	50.92±2.30	95.57±3.39	447.36±9.07	888.12±20.02
	EDGE (no LQC)	8.53±0.51	9.91±0.39	13.04±0.51	18.41±0.55	23.79±0.59	39.73±0.98	61.02±1.07	115.74±1.62	538.42±5.86	1055.96±14.25
	EDGE (LQC)	**2.17±0.17**	**2.51±0.12**	**3.99±0.17**	**5.95±0.19**	**8.02±0.19**	**14.25±0.41**	**22.28±0.55**	**43.87±1.14**	**200.36±4.80**	**411.48±14.24**

Table 5.2: QoS comparison: maximum delivery time for 95% of users with or without LQC (s)

System Load	Case	Message size (kB)									
		1	2	5	10	15	30	50	100	500	1000
0.04 msg/s	GPRS	1.81	3.24	7.59	14.26	21.09	39.52	60.56	114.41	509.97	940.16
	EDGE (no LQC)	3.17	4.99	10.61	19.64	28.29	50.80	79.65	148.03	632.71	1235.35
	EDGE (LQC)	**1.12**	**2.06**	**4.74**	**9.52**	**13.97**	**26.36**	**45.07**	**90.89**	**404.46**	**798.05**
0.1 msg/s	GPRS	26.11	26.42	31.22	37.42	43.64	60.93	86.84	147.42	631.6	1323.84
	EDGE (no LQC)	32.01	36.37	39.03	47.68	57.32	87.47	116.53	208.03	805.54	1529.37
	EDGE (LQC)	**4.22**	**5.87**	**11.54**	**18.43**	**25.68**	**47.28**	**73.31**	**141.31**	**605.64**	**1267.19**

Table 5.3: QoS comparison: first block delivery time with or without LQC (s)

Case	System Load (msg/s)				
	0.04	0.05	0.06	0.08	0.1
GPRS	1.24	1.97	2.75	4.41	5.87
EDGE (no LQC)	1.93	2.93	3.79	5.87	7.47
EDGE (LQC)	**0.28**	**0.47**	**0.67**	**1.18**	**1.69**

Table 5.4: QoS comparison: message discard rate with or without LQC (%)

Case	System Load (msg/s)				
	0.04	0.05	0.06	0.08	0.1
GPRS	4.04	8.40	13.49	25.50	36.53
EDGE (no LQC)	6.59	12.33	17.81	31.41	42.47
EDGE (LQC)	**0.53**	**1.58**	**3.3**	**7.77**	**13.44**

- **No reservation:** in this scheme,three channels are reserved for voice and four are shared between voice and data with preemptive priority to voice.
- **One channel reservation:** in this scheme, three channels are reserved for voice, one for data and three are shared between voice and data.
- **Two channels reservation:** in this scheme three channels are always reserved for voice, two for data and two are shared.

The other parameters are set up as follows: admission control is not used; the LQC algorithm is LA; all the users have the same priority.
The simulation results can be found in appendix B from Figures B.1 to B.23. It is clear that channel reservation has an important effect on all the measures of user perceived QoS. We portray the mean delivery time and the maximum delivery time for 95% of the users in Figures B.1 to B.20. Tables 5.5 and 5.6 show the delivery time of each message size at the load of 0.04 messages/s and of 0.1 messages/s and under the different data channel reservation policies. The best values are emphasized in bold in the tables.
When additional channels are reserved for data transmission, both the mean delivery time and the maximum delivery time for 95% of the users decrease. Such improvement is much more noticeable for large size messages. The more channels are reserved, the more improvement can be obtained. From the figures, we can also see a phenomenon. For the short messages, when one channel is reserved, the decrease of delivery time is apparent, especially the maximum delivery time for 95% of the users. However, when one more channel is reserved, such decrease is not so clear.
In Figures B.21 and B.22, we portray the same simulation results with respect to the message discard rate and the delivery time of first block. The numerial results are shown in tables 5.8 and 5.7 and the best results are portrayed in bold. As expected,

Table 5.5: QoS comparison: mean delivery time with different channel reservation policies (s)

System Load	Reserved data channels	Message size (kB)									
		1	2	5	10	15	30	50	100	500	1000
0.04 msg/s	0	1.56±0.14	2.05±0.11	3.31±0.13	5.73±0.22	7.73±0.28	13.37±0.60	20.42±0.78	37.1±1.53	166.32±6.37	323.29±16.57
	1	0.75±0.10	1.18±0.10	2.06±0.10	3.89±0.16	5.40±0.16	10.38±0.40	16.75±0.66	30.99±1.24	146.13±5.35	286.59±14.12
	2	**0.52±0.04**	**0.82±0.07**	**1.60±0.07**	**2.94±0.09**	**4.28±0.12**	**8.16±0.32**	**13.37±0.5**	**26.74±1.08**	**125.78±4.53**	**251.46±13.05**
0.1 msg/s	0	3.74±0.22	4.46±0.19	6.45±0.22	9.96±0.3	12.43±0.3	20.52±0.6	31.59±0.82	58.35±1.6	266.61±6.92	512.31±18.34
	1	2.69±0.18	3.38±0.15	5.03±0.19	7.54±0.22	10.27±0.25	17.64±0.51	28.4±0.73	51.77±1.37	237.00±5.90	470.39±15.92
	2	**2.17±0.17**	**2.51±0.12**	**3.99±0.17**	**5.95±0.19**	**8.02±0.19**	**14.25±0.41**	**22.28±0.55**	**43.87±1.14**	**200.36±4.80**	**411.48±14.24**

Table 5.6: QoS comparison: maximum delivery time for 95% of users with different channel reservation policies (s)

System Load	Reserved data channels	Message size (kB)									
		1	2	5	10	15	30	50	100	500	1000
0.04 msg/s	0	6.13	8.14	13.89	23.3	30.16	47.7	70.70	128.12	537.56	1006.01
	1	1.51	2.85	6.46	12.69	18.05	34.84	56.81	99.56	463.61	895.78
	2	**1.12**	**2.06**	**4.74**	**9.52**	**13.97**	**26.36**	**45.07**	**90.89**	**404.46**	**798.05**
0.1 msg/s	0	14.62	17.11	24.87	37.63	47.23	73.09	109.12	193.00	812.44	1596.16
	1	7.08	9.30	16.13	26.12	34.84	60.51	94.11	169.10	718.3	1368.65
	2	**4.22**	**5.87**	**11.54**	**18.43**	**25.68**	**47.28**	**73.31**	**141.31**	**605.64**	**1267.19**

the reservation of data channels reduces the discard rate and the first block delivery time.

Table 5.7: QoS comparison: first block delivery time with different channel reservation policies (s)

Reserved data channels	System Load (msg/s)				
	0.04	0.05	0.06	0.08	0.1
0	1.04	1.36	1.68	2.36	3.08
1	0.42	57	0.91	1.55	2.25
2	**0.28**	**0.47**	**0.67**	**1.18**	**1.69**

Table 5.8: QoS comparison: message discard rate with different channel reservation policies (%)

Reserved data channels	System Load (msg/s)				
	0.04	0.05	0.06	0.08	0.1
0	1.47	3.38	5.94	12.73	20.83
1	1.14	2.07	4.37	10.12	17.51
2	**0.53**	**1.58**	**3.3**	**7.77**	**13.44**

The data channel reservation can improve QoS offered to users at a cost of the increase of voice blocking rate. The voice blocking rate is shown in Figure B.23. As expected, the voice blocking rate increases when more channels are reserved for data transmission. In fact, the voice blocking rate is 6% for the no reservation scheme, 11% for one channel reservation and to almost 20% for 2 channels reservation (see table 5.9). If GSM network is already almost saturated, it would need a material cost to increase the total capacity of the network to be able to support a better quality of data service without decreasing the quality of voice call. It is necessary to find a compromise and our simulation can help to solve this dilemma by giving the variation of quality of service in a clear way.

Table 5.9: QoS comparison: Voice discarding rate (%)

Channels reserved for data	0	1	2
Voice Blocking rate	**5.78**	11.3	19.3

In the following sections, we will present simulation results to find the factors that influence the quality of service and the effects of the QoS policies.

5.8 Results for different LQC Algorithms

In this section, we will compare the results of the different LQC algorithms to see the influences on the user perceived QoS. The following two LQC algorithms are taken into account:

- **LA:** a pure Linking Adaptation is used according to the current channel conditions.
- **LA+IR:** the combination of LA and IR is used.

The other parameters used in this simulations are: two channels are reserved for data transmission; No admission control is considered; different classes of service are not taken into account to obtain the simulation results.

The simulation results can be found in the figures in appendix C. In Figures C.1 to C.10 we show the influence of using different LQC algorithms. The numerical results of the mean delivery time of each message size when using different LQC algorithms Some of the numerical results data are written in table 5.10, and the best values are marked in bold. We can see that compared with pure LA, the combination of LA and IR provides an advantage in mean delivery time. This advantage is lightly increased when we increase the system load. The results can verify the conclusion that we have made in section 4.2.4: the combination of LA and IR can improve the system performance. In Figures C.11 to C.20 we portray the same test but with respect to the maximum delivery time for 95% of the users. We can also see table 5.11 for some of the numerical results. And the part in bold indicates the best results. The advantage obtained by the combination of LA and IR is also clear.

Figures C.21 and C.22 show us the the comparison of message discard rate and system response time with different LQC algorithms. The numerical results can be found in tables 5.13 and table 5.12, in which the best results are portrayed in bold. As we have expected, the combination of LA and IR produces a less blocking rate and a lower delivery time of the first block. In brief, the simulation results show the advantage of the combination of LA and IR over the pure LA.

Table 5.14 gives a numerical analysis of the results obtained for mean delivery time 10kB messages and voice blocking rate using different channel reservation policies. The results in the table represent proportions compared to a base. The base chooses 100 as mean delivery time with 2 channels reserved, LA+IR and 0.1 message per second.

As a conclusion of the results with different LQC algorithms, we can say that the combination of LA and IR can improve the data quality of service without damaging the quality of voice call.

5.9 Results on the effects of admission control (AC)

In this section, we present the results to see if the admission control can influence the quality of service offered to the users.

When using the admission control, users with bad link condition will not be accepted

Table 5.10: QoS comparison: mean delivery time with different LQC algorithms (s)

System Load	Case	Message size (kB)									
		1	2	5	10	15	30	50	100	500	1000
0.04 msg/s	LA	0.52±0.04	0.82±0.07	1.60±0.07	2.94±0.09	4.28±0.12	8.16±0.32	13.37±0.5	26.74±1.08	125.78±4.53	251.46±13.05
	LA+IR	**0.38±0.04**	**0.55±0.02**	**1.13±0.05**	**2.05±0.06**	**2.91±0.07**	**5.72±0.19**	**9.44±0.29**	**17.68±0.60**	**87.11±2.69**	**171.74±7.46**
0.1 msg/s	LA	2.17±0.17	2.51±0.12	3.99±0.17	5.95±0.19	8.02±0.19	14.25±0.41	22.28±0.55	43.87±1.14	200.36±4.80	411.48±14.24
	LA+IR	**1.09±0.09**	**1.35±0.06**	**2.25±0.10**	**3.72±0.11**	**5.14±0.13**	**9.30±0.24**	**14.82±0.35**	**28.45±0.67**	**136.21±2.84**	**264.68±7.85**

Table 5.11: QoS comparison: maximum delivery time for 95% of users with different LQC algorithms (s)

System Load	Case	Message size (kB)									
		1	2	5	10	15	30	50	100	500	1000
0.04 msg/s	LA	1.12	2.06	4.74	9.52	13.97	26.36	45.07	90.89	404.46	798.05
	LA+IR	**0.81**	**1.40**	**3.18**	**6.29**	**8.90**	**17.97**	**29.01**	**52.78**	**265.28**	**544.06**
0.1 msg/s	LA	4.22	5.87	11.54	18.43	25.68	47.28	73.31	141.31	605.64	1267.19
	LA+IR	**1.59**	**2.92**	**6.38**	**11.73**	**16.59**	**31.44**	**51.45**	**98.88**	**438.23**	**857.72**

Table 5.12: QoS comparison: first block delivery time with different LQC algorithms (s)

Case	System Load (msg/s)				
	0.04	0.05	0.06	0.08	0.1
LA	0.28	0.47	0.67	1.18	1.69
LA+IR	**0.19**	**0.23**	**0.34**	**0.56**	**0.84**

Table 5.13: QoS comparison: message discard rate with different LQC algorithms (%)

Case	System Load (msg/s)				
	0.04	0.05	0.06	0.08	0.1
LA	0.53	1.58	3.3	7.77	13.44
LA+IR	**0.19**	**0.44**	**1.17**	**3.21**	**6.61**

Table 5.14: Influence of the number of reserved channels and the different LQC algorithms, for the mean delivery time of 10kB messages

Reserved channels	LQC algorithms		voice blocking rate
	LA	LA+IR	
0	258	179	5.78
1	207	137	11.3
2	151	100	19.3

by the system. Admission control cannot be used in the GPRS simulator mentioned in chapter 4 since it does take into account the varying channel condition. First of all, we will choose a threshold. If the link condition of a user (i.e. value of C/I) is lower than the threshold, it will be dropped even though the radio resources are available. In order to see the influences to the quality of service when different thresholds are used, we choose two thresholds for the simulations: $C/I = 5dB$ and $C/I = 10dB$. Under the current channel type (TU3), a threshold of 5dB means a block error rate of 30% when MCS1 is used and 10dB means a block error rate less than 10% when MCS1 is used. The other parameters are: two channels are reserved for data transmission; the LQC algorithm is pure LA; no other QoS policies are taken into account, which means the users have the same priority.

The figures of the results can be found appendix D. Some of the results are shown in tables from 5.15 to 5.18. In Figures D.1 to D.10, we portray the mean delivery time of each message size in different cases. Some of the numerical results are portrayed in table 5.15, in which the best results are marked in bold. We can see that the use of admission control decreases the mean delivery time. For the short messages, the improvement is not so clear (less than 1 second for 1kB message, see Figure D.1). However, this advantage is increased when the message size is increased. Note that the higher the threshold (10dB versus 5dB), the less the mean delivery time.

In Figures D.11 to D.20 we portray the same simulation results with respect to the maximum delivery time for 95% of the users. Some of the results can be found in table 5.16. We can have a lower maximum delivery time for 95% of the users due to the admission control. The higher the threshold, the more the advantage (see the bold part of table 5.16).

The same benefit can also be found in the delivery time for the first block (see Figure D.22 and table 5.17).

Finally, we can find the message discard rate in Figure D.21 and table 5.18. Notice the bold part in this table, when the system load is low, as expected, the use of

Table 5.15: QoS comparison: mean delivery time with or without Admission Control (s)

System Load	Case	Message size (kB)									
		1	2	5	10	15	30	50	100	500	1000
0.04 msg/s	Without Admission control	0.52±0.04	0.82±0.07	1.60±0.07	2.94±0.09	4.28±0.12	8.16±0.32	13.37±0.5	26.74±1.08	125.78±4.53	251.46±13.05
	Admission control (5dB)	0.37±0.02	0.59±0.03	1.20±0.05	2.25±0.1	3.23±0.07	6.03±0.17	10.22±0.32	19.57±0.62	93.55±2.68	185.26±7.13
	Admission Control (10dB)	**0.28±0.01**	**0.42±0.01**	**0.8±0.01**	**1.46±0.03**	**2.11±0.03**	**4.03±0.1**	**6.7±0.17**	**12.86±0.36**	**62.5±1.58**	**121.82±4.12**
0.1 msg/s	Without Admission Control	2.17±0.17	2.51±0.12	3.99±0.17	5.95±0.19	8.02±0.19	14.25±0.41	22.28±0.55	43.87±1.14	200.36±4.80	411.48±14.24
	Admission Control (5dB)	1.21±0.10	1.43±0.06	2.35±0.09	3.91±0.11	5.44±0.12	10.05±0.27	15.74±0.32	30.36±0.66	146.66±2.85	284.49±7.43
	Admission Control (10dB)	**0.50±0.04**	**0.65±0.03**	**1.22±0.04**	**2.13±0.05**	**3.05±0.07**	**5.59±0.11**	**9.07±0.17**	**18.12±0.57**	**85.16±1.55**	**168.57±3.9**

Table 5.16: QoS comparison: maximum delivery time for 95% of users with or without Admission Control (s)

System Load	Case	Message size (kB)									
		1	2	5	10	15	30	50	100	500	1000
0.04 msg/s	Without Admission Control	1.12	2.06	4.74	9.52	13.97	26.36	45.07	90.89	404.46	798.05
	Admission Control (5dB)	0.79	1.40	3.23	6.46	9.31	16.92	30.42	56.38	261.52	501.42
	Admission Control (10dB)	**0.57**	**0.9**	**3.23**	**3.79**	**5.48**	**10.29**	**17.18**	**32.64**	**145.37**	**291.28**
0.1 msg/s	Without Admission Control	4.22	5.87	11.54	18.43	25.68	47.28	73.31	141.31	605.64	1267.19
	Admission Control (5dB)	1.67	2.88	6.26	11.74	16.66	31.55	49.8	94.59	433.93	807.71
	Admission Control (10dB)	**0.75**	**1.27**	**2.86**	**5.42**	**8.04**	**15.25**	**24.77**	**49.17**	**217.52**	**411.16**

Table 5.17: QoS comparison: first block delivery time with or without Admission Control (s)

Case	System Load (msg/s)				
	0.04	0.05	0.06	0.08	0.1
Without Admission Control	0.28	0.47	0.67	1.18	1.69
Admission Control (5dB)	0.2	0.28	0.34	0.57	0.86
Admission Control (10dB)	**0.15**	**0.165**	**0.174**	**0.25**	**0.32**

admission control has the higher message discard rate than that without admission control. The higher the threshold, the higher the discard rate. As we know, without admission control, users with bad link condition will be accepted by the system if the resources are available and will occupy the radio resources for a long time (more radio blocks are needed) to finish their transmission. This resource occupation might not cause the extra discard rate if the traffic load is low. On the other hand, when admission control is used, users with bad channel condition will be dropped even though the radio resources might be available, which causes more users to be discarded. However, when the system load increases, the advantage of admission control appears. The refusal of users with bad channel condition will prevent the resources from being occupied for a long time, which reduce the possibility that a user being rejected due to the lack of resources and furthermore decrease the message discard rate. The cross point A and B in Figure D.21 are the changing points, where we can see the change of the performance with or without admission control. It is clear that the changing points are not the same if the thresholds are different.

In a word, the use of admission control does have an advantage on the data transfer time, but it might increase the message discard rate.

Table 5.18: QoS comparison: message discard rate with or without Admission Control (%)

Case	System Load (msg/s)						
	0.04	0.05	0.06	0.08	0.1	0.15	0.2
Without Admission Control	**0.53**	**1.58**	**3.3**	**7.77**	13.44	30.58	45.49
Admission Control (5dB)	6.11	6.57	6.71	8.79	11.87	23.91	36.36
Admission Control (10dB)	19.53	19.75	19.77	20.16	20.56	**23.71**	**29.07**

5.10 Results for QoS policy based on message size

In this section, we will compare the best effort policy with a QoS policy. With the best effort policy, all the users have the same priority and will be assigned with the same radio resources. On the other hand, we will use WFQ (Weighted Fair Queueing) scheme when the QoS policy is considered. We have assumed that in the WFQ scheme, the shorter messages will have priority over the long ones. In the following, a value of 3 is given to the weight of messages of 1KB, 2KB and 5KB; a weight of 2 is given to messages of 10KB up to 50KB and a weight of 1 is provided to message lengths of more than 100KB. The other simulation parameters are: LA is used as the LQC algorithm; admission control is not taken into account; two channels are reserved for data transmission.

Tables 5.19 and 5.20 show the mean delivery time and maximum delivery time for 95% of the users at the load of 0.04 messages/s and 0.1 messages/s. The best results are portrayed in bold in the tables. The figures of the results can be found in Figures E.1 to E.20.

From the results, we can see that the use of QoS classes provides an advantage in mean delivery time and maximum delivery time for 95% users for small and middle

Table 5.19: QoS comparison: mean delivery time with or without QoS policy (s)

System Load	Case	Message size (kB)									
		1	2	5	10	15	30	50	100	500	1000
0.04 msg/s	Best effort	0.52±0.04	0.82±0.07	1.60±0.07	2.94±0.09	4.28±0.12	8.16±0.32	13.37±0.5	26.74±1.08	125.78±4.53	251.46±13.05
	QoS policy	**0.46±0.05**	**0.69±0.05**	**1.34±0.06**	**2.58±0.08**	**3.73±0.13**	**7.11±0.25**	**11.39±0.39**	26.77±1.09	126.69±4.84	247.61±11.99
0.1 msg/s	Best effort	2.17±0.17	2.51±0.12	3.99±0.17	5.95±0.19	8.02±0.19	14.25±0.41	22.28±0.55	**43.87±1.14**	**200.36±4.80**	**411.48±14.24**
	QoS policy	**1.75±0.15**	**2.09±0.12**	**2.95±0.15**	**4.85±0.19**	**6.40±0.19**	**10.84±0.31**	**17.49±0.45**	44.67±1.16	209.68±5.18	414.39±14.41

Table 5.20: QoS comparison: maximum delivery time for 95% of users with or without QoS policy (s)

System Load	Case	Message size (kB)									
		1	2	5	10	15	30	50	100	500	1000
0.04 msg/s	Best effort	1.12	2.06	4.74	9.52	13.97	26.36	45.07	90.89	404.46	798.05
	QoS policy	**0.88**	**1.55**	**3.69**	**7.36**	**11.31**	**22.19**	**36.38**	**86.27**	**399.56**	**738.93**
0.1 msg/s	Best effort	4.22	5.87	11.54	18.43	25.68	47.28	73.31	**141.31**	**605.64**	1267.19
	QoS policy	**2.13**	**3.41**	**6.92**	**13.85**	**19.39**	**34.01**	**53.98**	143.45	646.95	**1252.28**

messages (<100kB). When the system load is low (0.04 messages/sec), this advantage is not obvious. The reason is the spectral efficiency realized by the high bit rate in an EDGE network. The radio resources are always sufficient for the users at low system load. When the system load increases, the advantage is increased. The WFQ scheme limits the resources used by large messages, which avoids the resources from monopolizing by the transmission of large messages. However, we can notice that there is a little performance decrease for large messages especially at high system load (e.g. Figures E.9 and E.19). The reason is: compared with the best effort policy, less resources are assigned to large messages when WFQ scheme is used for QoS classes, which results in this performance decrease. On the other hand, more resources are assigned to short messages so that they can finish their transmission more quickly, furthermore release the resources more quickly. These resources might be used by long messages. Therefore, the performance decrease for long message is not significant. As for the performance difference between the Figures E.9 and E.10 (the same phenomenon can be found between the Figures E.19 and E.20), the reason might be the lack of events for 1000kB messages. If we increase the traffic load, which means more events are produced, we can find that the difference disappears (see Figures E.23 and E.24, Figures E.25 and E.26). Tables 5.22 and 5.21 show the influence of QoS classes on message discard rate and first block delivery time and the bold part in the tables indicates the best results. The simulation results are also portrayed in Figures E.21 and E.22. In that case if the system load is low, the use of QoS classes cannot affect the discard rate and delivery time of first block. On the other hand, when the system load increases, both the message discard rate and the first block delivery time decrease due to the use of QoS classes.

Table 5.23 gives a numerical analysis of the results obtained for mean delivery time of 10kB messages and the voice blocking rate. The results in the table represent proportions compared to a base. The base chooses 100 as the mean delivery with 2

Table 5.21: QoS comparison: first block delivery time with or without QoS policy (s)

Case	System Load (msg/s)				
	0.04	0.05	0.06	0.08	0.1
Best effort	0.28	0.47	0.67	1.18	1.69
QoS policy	**0.28**	**0.42**	**0.58**	**1.03**	**1.56**

Table 5.22: QoS comparison: message discard rate with or without QoS policy (%)

Case	System Load (msg/s)				
	0.04	0.05	0.06	0.08	0.1
Best effort	0.53	1.58	3.3	7.77	13.44
QoS policy	**0.55**	**1.37**	**2.63**	**6.68**	**12.27**

channels reserved, classes of service according to the message size and 0.1 message per second.

To summarize these results we can say that the classes of service according to the message size are effective, without modifying the quality of service for the voice call.

5.11 Results for QoS policy based on differentiated users

In the last section, we presented the results of the QoS policy based on message size. Now, we want to investigate the system performance behavior when the QoS is provided to the differentiated users. A WFQ scheme is used so that different weights will be assigned to different user groups. However, the users in the same group have the same priority (same weight), regardless of their message size.

Table 5.23: Influence of the number of reserved channels and the classes of service based on the message size, for the mean delivery time of 10kB messages

Reserved channels	Classes of service		voice blocking rate
	Best effort	QoS	
0	188	165	5.78
1	152	132	11.3
2	111	100	19.3

The differentiation is based on the traffic classes defined in EDGE standard [29], which means that four traffic classes are taken into account in the simulation. We assume that all users have the same message length distribution but 10% of them will belong to class 1, 15% to class 2, 25% to class 3 and 50% to class 4. For each class of service, only three types of message lengths are considered so that we can get enough simulation events to obtain a more precise assessment. The message lengths are distributed as: **50% present 30kB, 35% 100kB, 15% 500kB**. The weight assignment is as follows: class 1 message weight is 4, class 2 is 3, class 3 is 2 and class 4 is 1. (see table 5.24)

Table 5.24: Proportion and weight of different classes of service

Classes of Service	Proportion (%)	Weight
Class 1	10	4
Class 2	15	3
Class 3	25	2
Class 4	50	1

The other parameters used for this simulation are: admission control is not used; the LQC algorithm is LA; two channels are reserved for data transmission. The users can be served using the different methods: differentiated queue, FCFS without priority and FCFS with priority. We will present the results to see if this policy can improve the QoS offered to the users; we also want to see with which queueing style the best quality of service can be achieved.

The QoS measures consist of the mean delivery time, the maximum delivery time for 95% of the users and the message discard rate.

We also want to make a performance comparison of the classes of service with the best effort users. The results based on the best effort policy come from another simulation during which all the users have the same priority and are served with best effort. The simulation parameters are identical to the ones mentioned in the previous part of this section.

The figures of the simulation results are portrayed in appendix F. In Figures F.1 to F.18 we portrayed the mean delivery time and maximum delivery time for 95% of the users of each class of service when using different methods. The numerical results can also be found in tables 5.25 and 5.26, but only the results at load of 0.1 messages/s are shown in these tables. In each table, there are four rows showing the results (separated with double-line). The first row indicates the delivery time for each message size without taking into account the classes of service (i.e. all the users are served with best effort); the second row shows the results per class when using the differentiated queue; the third row presents the results when the FCFS without priority is used; the fourth row points out the results when the FCFS with priority is used. The best values for each message size are portrayed in bold, whereas the worst are in italic.

Figures F.19 to F.21 portray the message discard rate. The numerical results are shown in table 5.27. There are four rows showing the results (separated with double-line). The first row indicates the global message discard rate at different system loads

Table 5.25: QoS comparison: mean delivery time of differentiated services (s), 0.1 messages/sec

Queueing Styles	Service Classes	Message size (kB)		
		30	100	500
	Best effort	29.81±0.28	72.57±0.7	319.87±4.91
Differentiated queue	Class 1	**12.94±0.50**	**35.33±0.85**	**152.0±4.10**
	Class 2	16.01±0.53	39.64±0.92	165.01±4.13
	Class 3	23.36±0.7	49.79±1.04	190.76±4.33
	Class 4	*49.4±1.13*	81.46±1.57	247.25±4.77
FCFS without priority	Class 1	20.98±0.53	45.27±1.02	178.87±6.08
	Class 2	22.80±0.44	49.69±0.91	204.8±5.89
	Class 3	26.24±0.38	60.66±0.88	261.88±5.65
	Class 4	35.34±0.33	*89.95±0.88*	*395.11±6.08*
FCFS with priority	Class 1	13.51±0.23	41.52±0.73	199.96±5.3
	Class 2	21.76±0.36	51.69±0.88	226.68±5.92
	Class 3	25.27±0.32	64.11±0.86	282.41±5.78
	Class 4	32.59±0.27	71.43±0.65	221.61±4.83

Table 5.26: QoS comparison: maximum delivery time for 95% of users of differentiated services (s), 0.1 messages/sec

Queueing Styles	Service Classes	Message size (kB)		
		30	100	500
	Best effort	94.15	199.12	855.85
Differentiated queue	Class 1	**37.24**	**107.45**	**428.75**
	Class 2	47.33	121.77	464.16
	Class 3	79.15	167.13	564.09
	Class 4	*208.75*	*286.58*	743.95
FCFS without priority	Class 1	69.22	132.88	488.69
	Class 2	72.48	148.02	557.86
	Class 3	84.91	176.23	708.14
	Class 4	108.93	257.95	*1085.9*
FCFS with priority	Class 1	43.27	123.86	557.83
	Class 2	69.43	152.21	642.7
	Class 3	79.53	187.17	795.98
	Class 4	97.28	195.44	435.22

without taking into account the classes of service (i.e. best effort); the second row shows the results per class when using the differentiated queue; the third row presents the results when the FCFS without priority is used; the fourth row points out the results when the FCFS with priority is used. Again, the best values for each system load are marked in bold, whereas the worst are in italic.

Table 5.27: QoS comparison: message discard rate of differentiated services (%), 0.1 messages/sec

Queueing Style	Service Classes	System Load (msg/s)				
		0.04	0.05	0.06	0.08	0.1
	Best effort	11.3	19.96	28.02	41.84	53.31
Differentiated Queue	Class 1	6.70	9.67	13.24	19.02	24.84
	Class 2	10.82	14.73	18.91	27.78	34.23
	Class 3	16.84	23.71	29.19	40.71	48.73
	Class 4	*30.31*	*40.42*	*48.6*	*61.24*	*70.13*
FCFS without priority	Class 1	11.52	19.73	28.30	42.42	52.80
	Class 2	11.41	18.45	28.35	42.0	52.55
	Class 3	11.67	19.08	28.44	42.38	53.25
	Class 4	11.68	19.38	28.34	42.56	52.77
FCFS with priority	Class 1	**2.18**	**3.24**	**5.43**	**9.65**	**13.71**
	Class 2	9.0	15.22	22.25	34.98	44.41
	Class 3	9.13	15.18	22.41	34.73	44.53
	Class 4	12.19	19.76	28.77	43.70	54.73

First of all, let us assess the delay difference between the users with different classes and with the case in which all users operate under a best-effort policy.

- **The delivery time**

  As expected, no matter which method is used, both the mean delivery time and maximum delivery time for 95% of the users in class 1 are the lowest, whereas the delivery time of the users in class 4 are the highest since we assigned more transfer turns to users with higher priority. However, there are the exceptions: when FCFS with priority is used, it can be seen that the mean delivery time of users in class 4 is even lower than users in class 2 (see Figure F.9); and the maximum delivery time for 95% of the users decreases as the load increases (see Figure F.18). It can be explained as follows:

  As we have mentioned in subsection 4.2.7, users in class 1 can occupy the resources used by users in class 4 if there are already 4 messages being transferred. In that case if there are more than one class 4 messages being transferred, the one with longest queue will be replaced, which means the large size messages with bad channel condition have much more opportunity to be discarded. Therefore, only a few large size messages will suffer the long transfer time, which leads to a lower mean and max transfer time. As the load increased, more and more large size messages with bad channel condition are discarded, which results in the improvement of the performance.

  When compared with the users under a best effort policy, the mean transfer time for the users of first three classes are lower than best effort users, whereas the performance of class 4 users is worse than best effort users except for the cases mentioned above. Note that in Figures F.7 and F.16 the delivery time of class 4 users is lower than best effort users. The reason might be that: when a differentiated queue is used, less messages could be transferred simultaneously. Therefore, more resources could be assigned to class 4 users.

  We can see that the best values (bold in the tables) appear when differentiated queue is used. It seems that the differentiated queue performs better than the other two methods. We also find that most of the worst results (italic part in the tables) appear when the FCFS without priority is used. However, we are only interested in the quality of service of the first two classes, and later on, we will make a comparison of these three methods.

- **Message discard rate**

  The message discard rate of each classes of services depends on the method that they used.

  - **Differentiated queue**

    It is clear that the discard rate of class 1 users is the lowest, whereas the discard rate of class 4 users is the highest. Compared with best effort users, the discard rate of the first two classes is lower than best effort users since most of the resources are allocated for them to finish their transmission. As expected, the discard rate of class 4 users are higher than best effort users. With respect to the class 3 users: their discard rate is higher than best effort users at low system load and is lower than best effort users at high system load (see Figure F.19).

  - **FCFS without priority**

    In Figure F.20, no matter which classes are concerned, there is no difference in message discard rate. The reason is that the utilization of the queues follows the policy of first-come-first-use, which means that all the users have the same probability to be discarded due to the fullness of the queues.

  - **FCFS with priority** It can be seen from Figure F.21 that the discard rate of class 1 users is much lower than other users because they can occupy the resources being used by class 4 users when they arrive. There is no difference between the discard rate of class 2 and class 3 users. The performance of these three classes of users is better than the one for best effort users. As expected, class 4 users suffer the highest discard rate.

Regarding the overall best and worst results: the best results are obtained when the FCFS with priority is used; whereas the worst results appear to class 4 users when the differentiated queue is used. The reason is: when FCFS with priority is used, class 1 users can occupy the resources being used by class 4 users, which decrease the message block rate of class 1 users. When the differentiated queue is used, the users of the same class cannot be simultaneously transferred. This method might give benefit to the class with low population (e.g. class 1). However, it brings disadvantage to the class with high population (e.g. class 4) because they cannot occupy the places belonging to other classes. Thus, much more of the transfer requirements are refused even if the resources are available.

We have seen that the performances would not be the same if different methods are used. We portray the comparison of these three methods in Figures F.22 to F.35. Only the measures of class 1 and class 2 users are compared because we want to find the best method to guarantee the QoS requirements of the first two classes. With respect to the delivery time, it can be seen that the differentiated queue performs well for most of the cases. However, there are some exceptions. In Figures F.22 and F.23, we can see that the FCFS with priority performs well when system load is low. But the difference is not so significant. On the other hand, when the system load is high, the differentiated queue performs much better than the other two methods. As regards the message discard rate, the FCFS with priority performs the best for class 1 users (see Figure F.34), whereas the differentiated queue performs well for class 2 users (see Figure F.35). In general, both the differentiated queue and the FCFS with priority have their own advantage. We have to make a compromise to meet the most important QoS requirement. For instance, differentiated queues should be used if we want a lower transfer delay. On the other hand, if we want to decrease the message discard rate of class 1 users, we could choose FCFS with priority.

It could be seen that in some of the curves presented in this section, it is possible to observe a "performance improvement" as the load increases (see Figure F.18). This improvement is not the case that we explained above. The improbable behavior can be explained by the confidence interval of the values portrayed. In fact, even if the lower bound of such an interval increases with the load, it is possible that the individual points decrease in value, thus giving the false impression that the performance can be improved with increasing load. Figure F.36 portrays the maximum delivery time for 95% of the users for 500kB messages. The confidence intervals are portrayed for each curve. Figure F.37 portrays the lower bound of each confidence interval, which illustrates such a behavior.

Contrary to the classes of service based on message size which can degrade the performances of the longest messages, this policy involves a more significant deterioration of the performances of those background services. This policy does not require to make the sorting of the messages according to their length, a user using the same class of service for all its messages. Table 5.28 gives a numerical analysis of the results obtained for mean delivery time of 10kB messages and the voice blocking rate. The results in the table represent proportions compared to a base. The base chooses 100 as the mean delivery with 2 channels reserved, classes of service according to the differentiated users using differentiated queue and 0.1 message per second.

As a conclusion of the results with classes of service based on differentiated users, we can say that this policy of classes of service makes it easier to classify a user group. The QoS improvements of high priority users are achieved by damaging the QoS of background service, without affecting the QoS of voice call. As for the different methods, the differentiated queue and the FCFS with priority have their own advantage and a compromise need to be made to meet the most important QoS requirements.

Table 5.28: Influence of the number of reserved channels and the classes of service based on the differentiated users, for the mean delivery time of 30kB messages

Queueing styles	Reserved channels	Classes of service					voice blocking rate
		Best effort	Class 1	Class 2	Class 3	Class 4	
Differentiated queue	0	298	152	183	267	502	5.78
	1	280	134	160	229	464	11.3
	2	262	100	146	197	437	19.3
FCFS without priority	0	298	224	244	273	364	5.78
	1	280	194	220	245	328	11.3
	2	262	176	196	219	296	19.3
FCFS with priority	0	298	154	232	268	326	5.78
	1	280	130	209	243	302	11.3
	2	262	111	181	204	269	19.3

Chapter 6

Conclusion

When deploying an EDGE network, it is important to be able to evaluate the system performance and the quality of service (QoS). The objective of our work was to propose a tool to estimate user perceived QoS in an EDGE network. First of all, we studied the functionality of EDGE with the help of the ETSI standards. This investigation allowed us to figure out the aspects of EDGE which impact the quality of service offered to the users. Then we conducted a literature review. We found that in the literature, many simulation models had been proposed, but most of them focus on the global system performance, such as system throughput, channel utilization, etc. There are also analytical models but they have their own limitations. Some of them cannot manage the multi-channel communications; some others have limitations on the classes of services. Besides, the impacts of voice calls are not taken into account. Therefore, we concluded that the analytical models cannot achieve our goal of assessing the user-perceived QoS and proposed a new type of simulator.

A GPRS simulator has been proposed in [32] to assess user QoS in a GPRS network. In this simulator, the RLC protocol, classes of services and resource allocation were modelled. The QoS measures included *the delivery time of a message*, *the delivery time for the first block of data* and *the blocking rate.* All of these features were used to estimate the QoS offered to the users in an EDGE network. However, in order to give an accurate model to simulate an EDGE system, we needed to make some significant modifications. The key tasks to achieve this were the estimation of the channel conditions and the use of Link Quality Control (LQC). Therefore, we added a channel condition generator to simulate the varying channel conditions and propose a simple algorithm to implement link quality control (see subsection 4.2.4). Besides, some modifications were also carried out in packet segmentation and queueing system.

With the simulator, we have produced a series of results to assess the factors that influence the QoS offered to the users.

The importance of link quality control was firstly verified: the use of link quality control can significantly improve user perceived QoS. As for the data channel reservation, the number of channels reserved for data has a great influence on the quality of service. Our simulator makes it possible to evaluate the improvement of the quality of service resulting from an increase in the number of channels reserved to the data transmission, and, at the same time, quantify the degradation of the performance of voice calls. Our tool, thus, gives the operators an invaluable help to make management decisions. Besides, we also proposed other QoS policies to improve the quality of service offered to users. These policies consist of the use of effective LQC algorithm, the use of admission control and the classes of service.

Thanks to the benefit of code combining, the combination of LA and IR can achieve better data quality of service without affecting the quality of service of voice calls. Admission control is another technology that can decrease the data transfer time, but it might cause relatively high message discard rate at low system load.

As for the classes of service, in general, we can say that the introduction of classes of services allows a better exploitation of the network and takes better into account the customer requirements. We investigated two different policies to manage the classes of service (classes of service based on message size and classes of services based on differentiated users). With the QoS policy based on message size, the quality of service for the data transmission can be improved, without damaging the quality of service of the voice call. With respect to the classes of service by user, unlike the classes of service based on messages size where the messages need to be sorted according to their length, a user uses the same class of service for all its messages, regardless of their length. This makes it easier to classify the user groups. Based on the characteristic of the queueing system in our simulator, we proposed three methods to serve the differentiated services: differentiated queue, FCFS without priority and FCFS with priority. No matter which method was used, this QoS policy improves the quality of service of high priority users by significantly sacrificing the performances of those background services. We also made a comparison of these methods in order to find the best method for the quality of service of the first two classes. We found that both the differentiated queue and the FCFS with priority give benefit to the first two classes. Therefore, we could use either of the method according to the quality of service that we want to guarantee.

In brief, with these simulations, we can draw the following conclusions:

- It is important to use link quality control in an EDGE network to improve the QoS seen by the user. Regarding the different LQC algorithms, the combination of LA and IR can achieve a better quality of service than pure LA.

- The data channel reservation is another important factor to influence the QoS seen by the user. It can greatly improve the QoS seen by the user, whereas it has a negative impact on voice blocking.

- The use of admission control does have an important benefit effect on the data transfer time at a cost of high message discard rate at low system load. However, when the system load is high, admission control can significantly improve the QoS seen by the user.

- The use of QoS classes can improve the delivery time for short and middle size messages without significantly affecting the performance of large size messages.

- As for QoS policy based on differentiated queue, the high priority users have a QoS improvement and the class 4 users suffer the worst quality of services. Based on the simulation results, either the differentiated queue or the FCFS with priority has their advantage. Which methods can be used depends on the QoS that we want to guarantee.

Finally, as regards further work, the following aspect would be interesting:

- The multiple cells model should be developed. At the moment, the simulator only takes a single cell into account. In such a case, not only the channel selection, but also the cell selection should be modelled in the simulator. Besides, the resource allocation should be careful of not assigning the channels of different cells to one user.

- When the multiple cells are considered, the handoff of voice calls and data should also be studied because users do not want their calls or data transfers interrupted due to the switching of the cell. We need to find a mechanism (for example, channel reservation) to make sure that handoff calls will not be rejected due to the lack of resources.

- Another interesting work would be to assess real-time IP services such as voice or video over IP. In order to implement that assessment, we should consider the traffic model, the cell and channel selection and the channel allocation to satisfy the QoS requirement. In addition to the transfer delay, jitter should also be taken into account as a measure of QoS.

Bibliography

[1] The 3rd Generation Partnership Project(3GPP), http://www.3gpp.org/. *3GPP TSG RAN Meeting 21.*, 2003.

[2] A.Furuskar, J.Nalund, and H.Olofsson. EDGE: Enhanced data rates for GSM and TDMA/136 evolution. *Ericsson Review*, January 1999.

[3] A.Furuskar, M.Frodigh, H.Olofsson, and J.Skod. System performance of EDGE, a proposal for enhanced data rates in existing digital cellular systems. In *IEEE VTC 1998*, volume 2, pages 1284–1289, 1998.

[4] A.Furuskar, M.Frodigh, H.Olofsson, and J.Skod. Capacity evaluation of the EDGE concept for enhanced data rates in GSM and TDMA/136. In *IEEE VTC 1999*, volume 2, pages 1648–1652, 1999.

[5] B.H.Walke. *Mobile Radio Networks: Networking, Protocols and Traffic Performance.* New York: Wiley, 2002.

[6] B.Mehta and J.smith. Performance analysis of link adaptation in wireless data networks. In *IEEE GLOBECOM'00*, volume 3, pages 1422–1426, December 2000.

[7] D. Chase. Code combining - a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets. In *IEEE TRANSACTIONS ON COMMUNICATIONS*, volume 33, pages 385–393, 1985.

[8] C.Johansson, L.de Verdier, and F.Khan. Performance of different scheduling strategies in a packet radio system. In *Universal Personal Communications*, volume 1, pages 267–271, 1998.

[9] D.Bertsekas and R.Gallagher. *Data Networks.* Prentice Hall, 1987.

[10] D.Molkdar and S.Lambotharan. Impact of header error on EGPRS link level performance in IR mode. *3G Mobile Communication Technologies*, pages 114–118, March 2001.

[11] D.Molkdar and S.Lambotharan. Link performance evaluation of EGPRS in LA and IR modes. In *IEEE VTC 2000*, pages 1498–1506, September 2000.

[12] D.Molkdar, W.Featherstone, and S.Lambotharan. An overview of EGPRS: the packet data component of EDGE. *Electronics and Communication Engineering Journal*, 14:21–38, 2002.

[13] D.Soldani, N.Lokuge, and A.Kuurne. Service performance monitoring for EGPRS networks based on treatment classes. *IWQOS 2004*, pages 121–128, 2004.

[14] D.Todinca, P.Perry, and J.Murphy. Algorithms for resource allocation in data transfer over EGPRS networks. In *ECUMN 2002*, pages 246–250, April 2002.

[15] Ericsson White Paper, http://www.ericsson.com. *EDGE: Introduction of high-speed data in GSM/GPRS networks.*, 2002.

[16] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 05.01: Digital cellular telecommunications system (Phase 2+); Physical layer on the radio path; General description*, version 8.6.3 edition, 2000.

[17] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 05.03: Digital cellular telecommunications system (Phase 2+); Channel coding*, version 8.6.0 edition, 2000.

[18] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 03.64: Digital cellular telecommunications system (Phase 2+); Overall description of the General Packet Radio Service (GPRS) Radio interface; Stage 2*, version 8.9.0 edition, 2001.

[19] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 04.60: Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Mobile Station (MS) - Base Station System (BSS) interface; Radio Link Control/Medium Access Control (RLC/MAC) protocol*, version 8.5.0 edition, 2001.

[20] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 04.64: Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Mobile Station - Serving GPRS Support Node (MS-SGSN) Logical Link Control (LLC) layer specification*, version 8.7.0 edition, 2001.

[21] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 05.02: Digital cellular telecommunications system (Phase 2+); Physical layer on the radio path; Multiplexing and multiple access on the radio path*, version 8.9.0 edition, 2001.

[22] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 05.04: Digital cellular telecommunications system (Phase 2+);Modulation*, version 8.4.0 edition, 2001.

[23] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 05.05: Digital cellular telecommunications system (Phase 2+); Radio transmission and reception*, version 8.9.0 edition, 2001.

[24] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 08.18: Digital cellular telecommunications system (Phase2+); General Packet Radio Service (GPRS); Base Station System (BSS) - Serving GPRS Support Node (SGSN); BSS GPRS Protocol (BSSGP)*, version 8.9.0 edition, 2001.

[25] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 03.60: Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS) Service description; Stage 2*, version 7.9.0 edition, 2002.

[26] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 08.52: Digital cellular telecommunications system (Phase 2+); Base Station Controller - Base Tranceiver Station (BSC-BTS) interface; Interface principles*, version 8.0.1 edition, 2002.

[27] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 23.060: Universal Mobile Telecommunications System (UMTS);Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Service description; Stage 2*, version 5.2.0 edition, 2002.

[28] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 02.60: Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Service description; Stage 1*, version 5.3.0 edition, 2003.

[29] European Telecommunications Standards Institute (ETSI), http://www.etsi.org/. *GSM 23.107: Universal Mobile Telecommunications System (UMTS); Quality of Service (QoS) concept and architecture*, version 5.12 edition, 2004.

[30] European Telecommunications Standards Institute(ETSI), http://www.etsi.org/. *GSM 04.65: Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Mobile Station (MS) - Serving GPRS Support Node (SGSN); Subnetwork Dependent Convergence Protocol (SNDCP)*, version 8.2.0 edition, 2001.

[31] F.Bodevin. Mesure de la qualité de service usager d'un réseau GPRS. Master's thesis, École Polytechnique de Montréal, 2002.

[32] F.Bodevin and B.Sansò. User quality of service for GPRS networks. *Cahier du GERAD*, November 2001.

[33] Internet, http://www.rfc-editor.org/rfc/rfc1144.txt. *RFC 1144*, february 1990 edition, 1990.

[34] J.Cai, L.F.Chang, K.Chawia, and X.Qiu. Providing differentiated services in EGPRS through packet scheduling. *IEEE Globecom*, 3:1515–1521, 2000.

[35] K.Murota and K.Hirade. GMSK modulation for digital mobile radio telephony. *IEEE Transactions on Communications*, 29, July 1981.

[36] L.F.Chang, X.Qiu, K.Chawia, and J.Cai. Providing differentiated services in EGPRS through radio resource management. *IEEE ICC*, 7:2296–2301, 2001.

[37] L.Zhao, J.W.Mark, and Y.C.Yoon. A combined link adaptation and incremental redundancy protocol for enhanced data transmission. In *IEEE GLOBECOM '01*, volume 2, pages 1277–1281, 2001.

[38] M.E.R.Angeles, D.L.Rodriguez, and F.A.C.Pérez. A new EDGE medium access control mechanism using adaptive traffic load slotted aloha. In *IEEE VTC 2001*, volume 3, pages 1358–1362, 2001.

[39] M.Mzyece and J.Dunlop. Performance of EGPRS incremental redundancy in interference-limited environments. *Wireless Personal Multimedia Communications*, 2:873–877, October 2002.

[40] M.Salmenkaita, J.Gimenez, and P.Tapia. Optimizing the GSM/EDGE air interface for multiple services with dynamic frequency and channel assignment. In *IEEE VTC 2002 FALL*, volume 4, pages 2215–2219, 2002.

[41] M.Zhang and A.MacDonald. Performance simulation for shared voice/data EDGE channel. In *IEEE VTC 2001*, volume 3, pages 1398–1401, 2001.

[42] N.K.Shankaranarayanan, Z.Jiang, and P.Mishra. User-perceived performance of web-browsing and interactive data applications in TDMA packet wireless networks. *Multiaccess, Mobility and Teletraffic for Wireless Communications*, 5:73–84, 2000.

[43] N.K.Shankaranarayanan, Z.Jiang, and P.Mishra. Performance of a shared packet wireless network with interactive data users. *Mobile Networks and Applications*, 8:279–293, 2003.

[44] P.Schramm, H.Andreasson, C.Edholm, N.Edvardsson, M.Hook, S.Javerbring, F.Muller, and J.Skod. Radio interface performance of EDGE, a proposal for enhanced data rates in existing digital cellular systems. In *IEEE VTC 1998*, volume 2, pages 1064–1068, 1998.

[45] P.Stuckmann and J.Franke. Performance characteristics of the enhanced general packet radio service for the mobile internet access. In *Second International Conference on 3G Mobile Communication Technologies, London, UK*, 2001.

[46] S.Eriksson, A.Furuskar, M.Hook, S. Javerbring, H.Olofsson, and J.Skold. Comparison of link quality control strategies for packet data services in EDGE. In *IEEE VTC 1999*, volume 2, pages 938–942, 1999.

[47] S.Kallel. Analysis of a type II hybrid ARQ scheme with code combining. In *IEEE TRANSACTIONS ON COMMUNICATIONS*, volume 38, pages 1133–1137, 1990.

[48] S.Nikkarinen and M.Moisio. Network level performance evaluation of EGPRS. In *IEEE VTC 2002 FALL*, volume 3, pages 1811–1815, 2002.

[49] X.Lagrange. CIR cumulative distribution in a regular cellular network. In *Rapport interne ENST 2000D001*, 2000.

[50] X.Qiu, J.Chuang, K.Chawla, and J.Whitehead. Performance comparison of link adaptation and incremental redundancy in wireless data networks. In *IEEE WCNC*, volume 2, pages 771–775, September 1999.

[51] Y.Cao, H.R.Sun, and K.S.Trivedi. Performance analysis of reservation media-access protocol with access and serving queues under bursty traffic in GPRS/EGPRS. In *IEEE VTC 2003*, volume 52, pages 1627–1641, 2003.

[52] Y.C.B.Silva, W.M.de Sousa, F.R.P.Cavalcanti, and T.F.Maciel. Capacity and qos enhancement for data transmission in EGPRS through packet scheduling and smart antennas. In *IEEE ITS 2002*, September 2002.

Appendix A

Results for effects of Link Quality Control

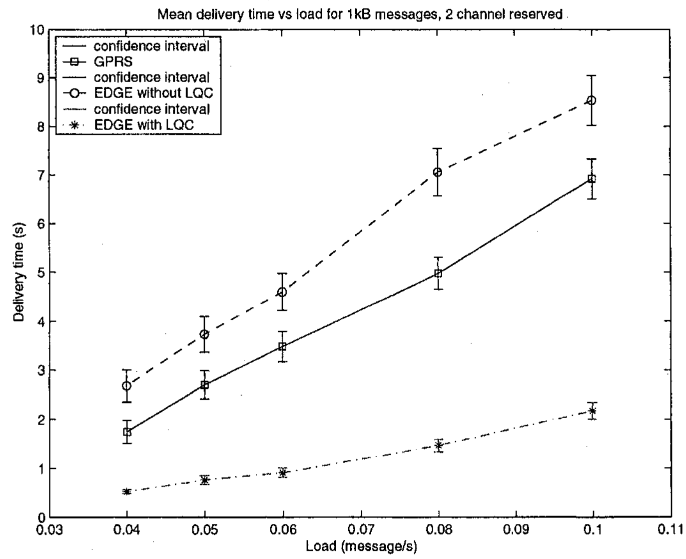


Figure A.1: Mean delivery time, 1KB messages, 0.06 call/s, call length 60 s

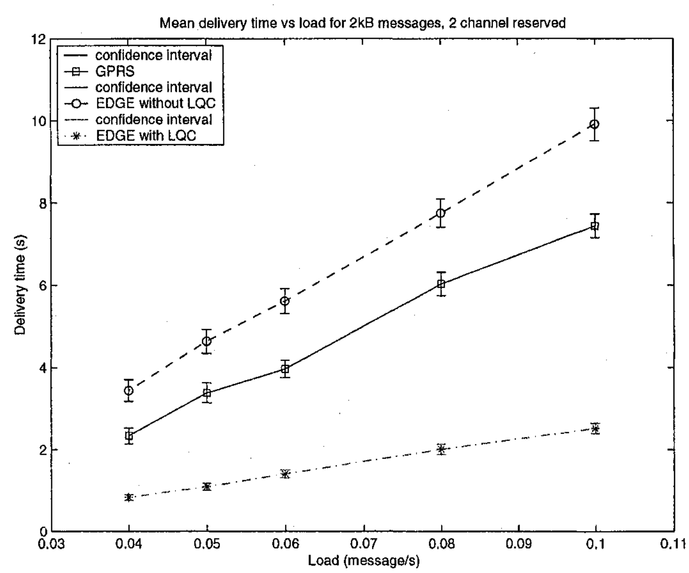


Figure A.2: Mean delivery time, 2KB messages, 0.06 call/s, call length 60 s

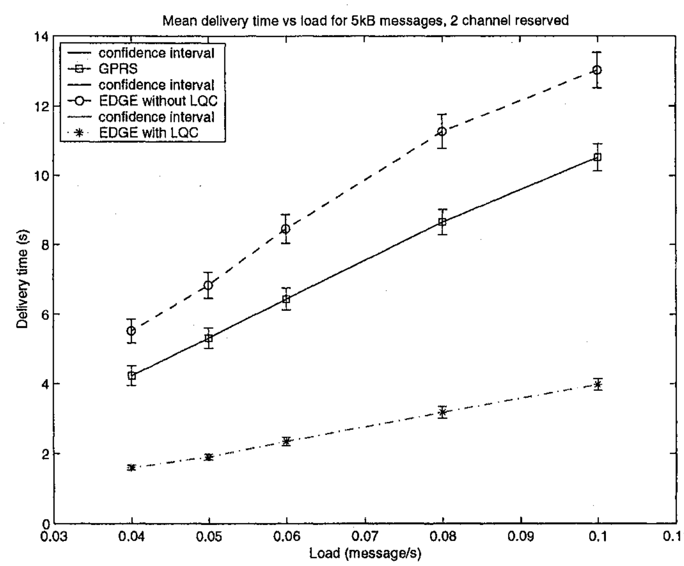


Figure A.3: Mean delivery time, 5KB messages, 0.06 call/s, call length 60 s

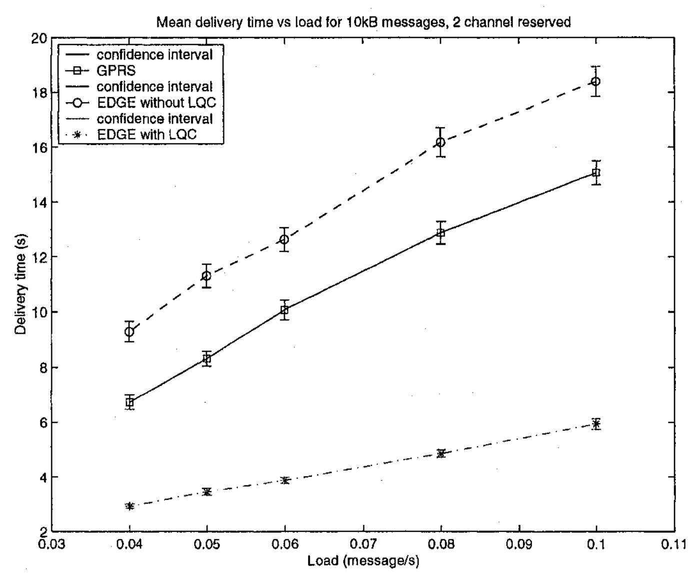


Figure A.4: Mean delivery time, 10KB messages, 0.06 call/s, call length 60 s

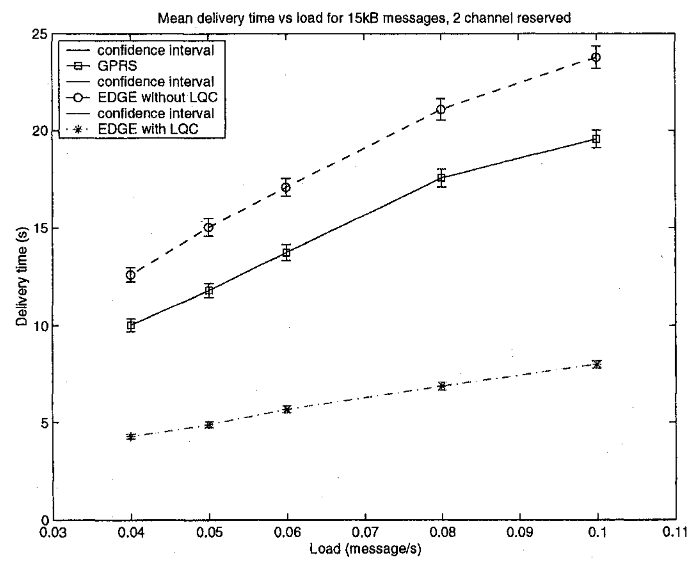


Figure A.5: Mean delivery time, 15KB messages, 0.06 call/s, call length 60 s

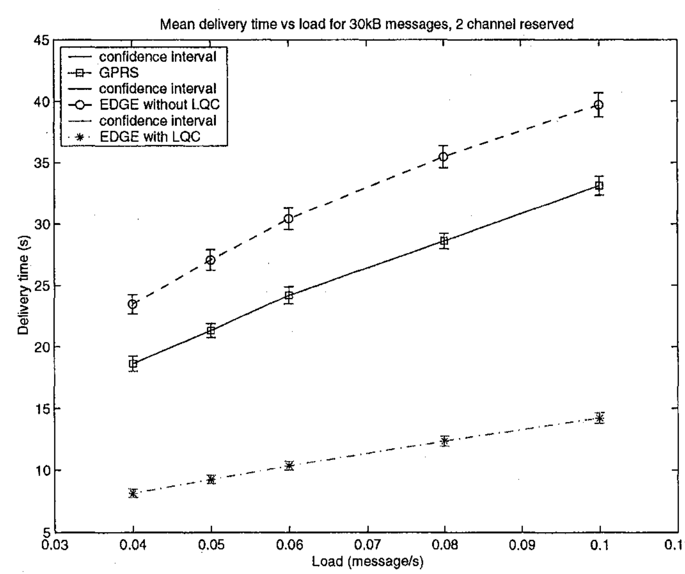


Figure A.6: Mean delivery time, 30KB messages, 0.06 call/s, call length 60 s

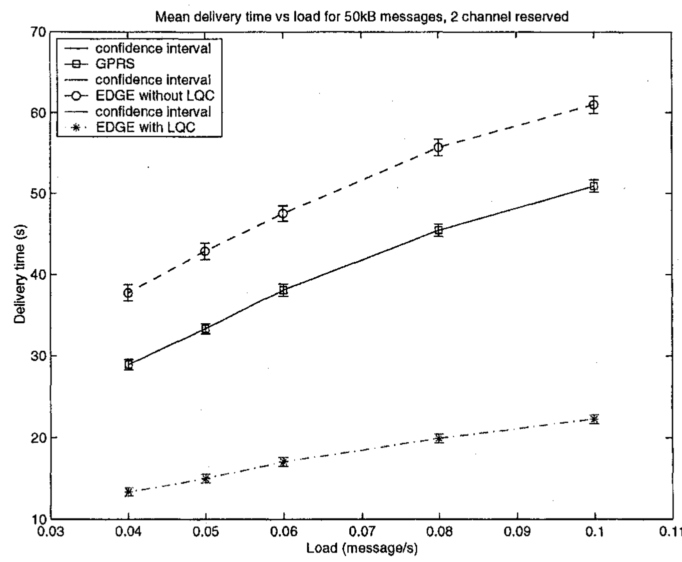


Figure A.7: Mean delivery time, 50KB messages, 0.06 call/s, call length 60 s

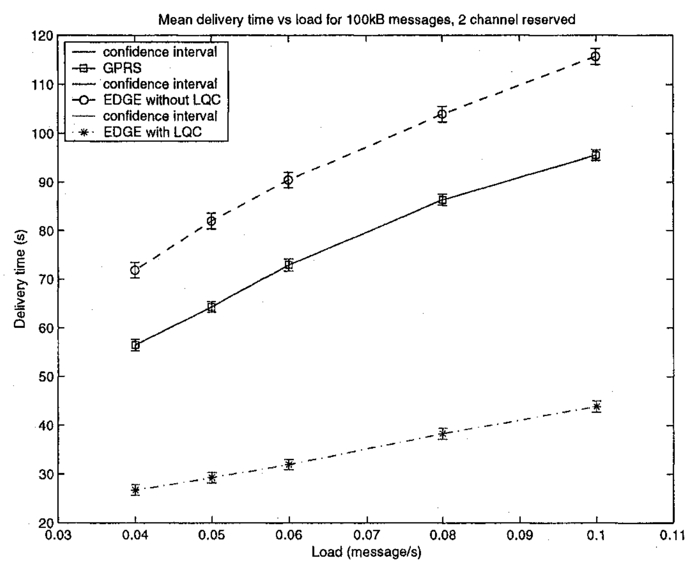


Figure A.8: Mean delivery time, 100KB messages, 0.06 call/s, call length 60 s

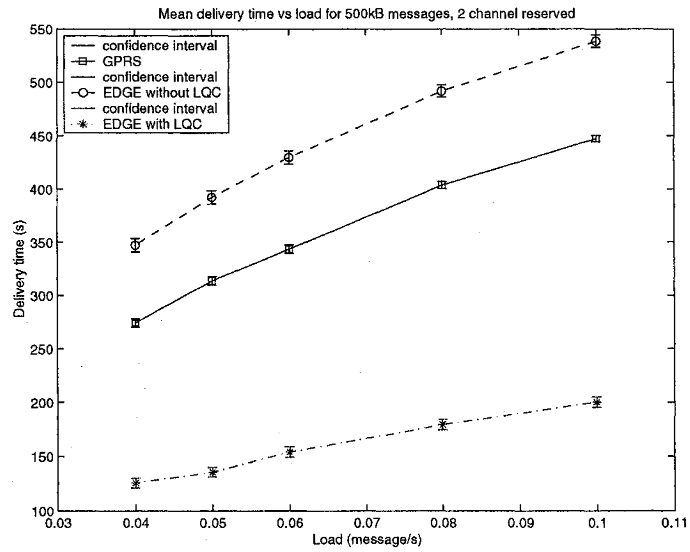


Figure A.9: Mean delivery time, 500KB messages, 0.06 call/s, call length 60 s

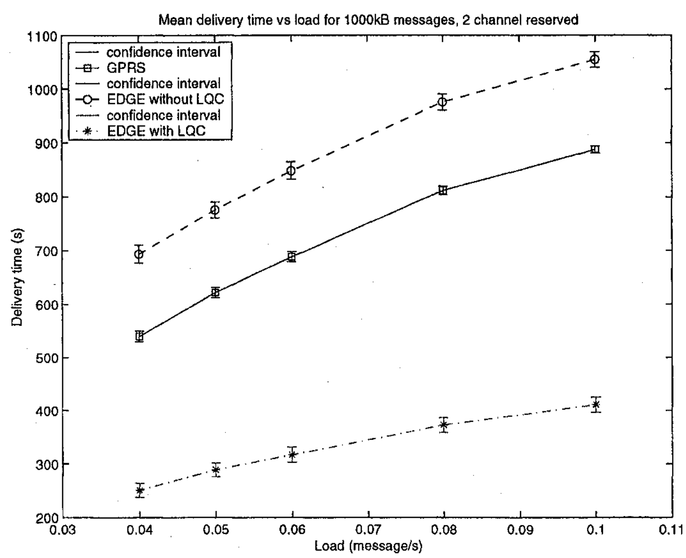


Figure A.10: Mean delivery time, 1000KB messages, 0.06 call/s, call length 60 s

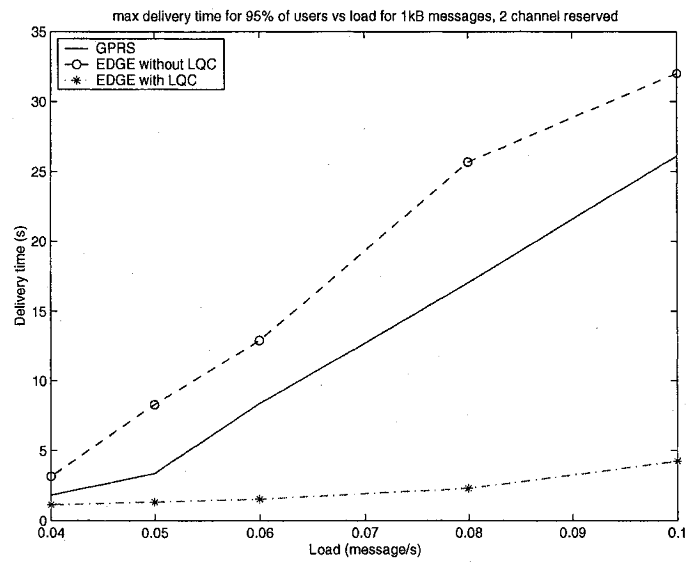


Figure A.11: Maximum delivery time for 95% of users, 1KB messages, 0.06 call/s, call length 60 s

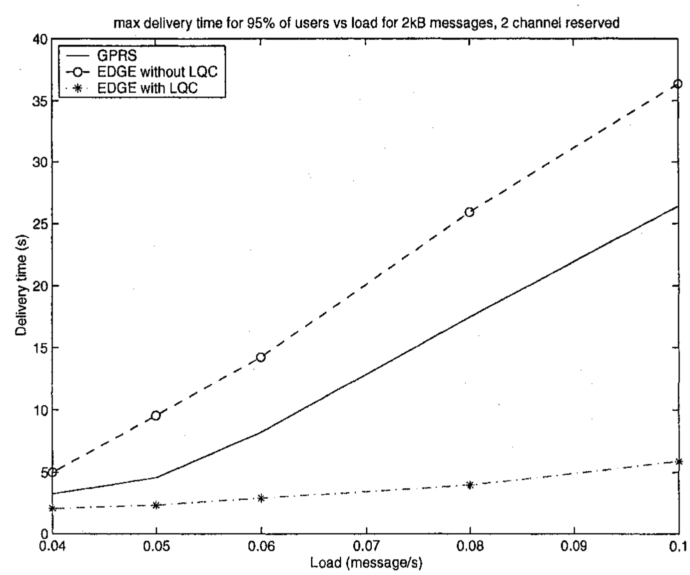


Figure A.12: Maximum delivery time for 95% of users, 2KB messages, 0.06 call/s, call length 60 s

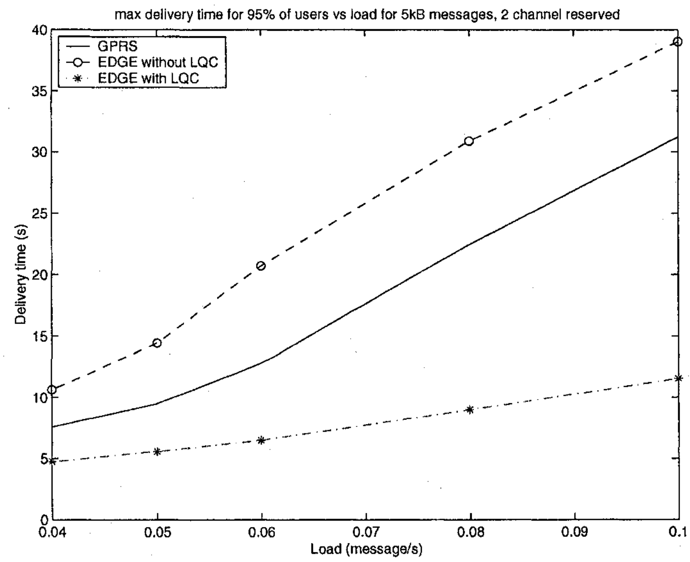


Figure A.13: Maximum delivery time for 95% of users, 5KB messages, 0.06 call/s, call length 60 s

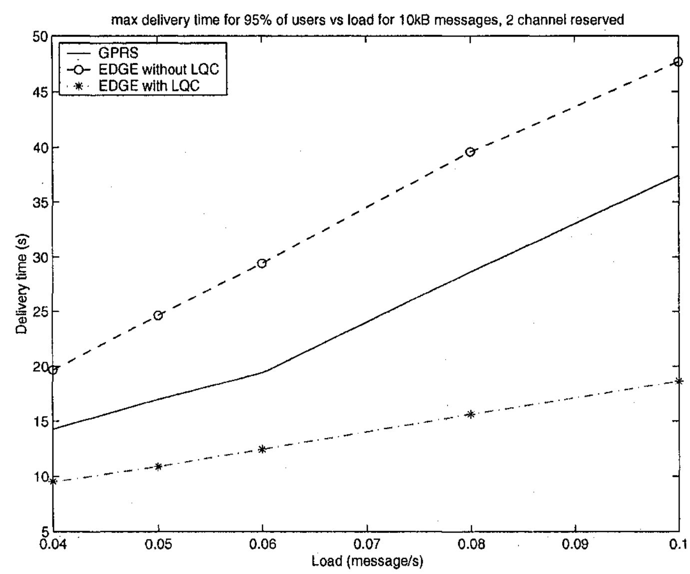


Figure A.14: Maximum delivery time for 95% of users, 10KB messages, 0.06 call/s, call length 60 s

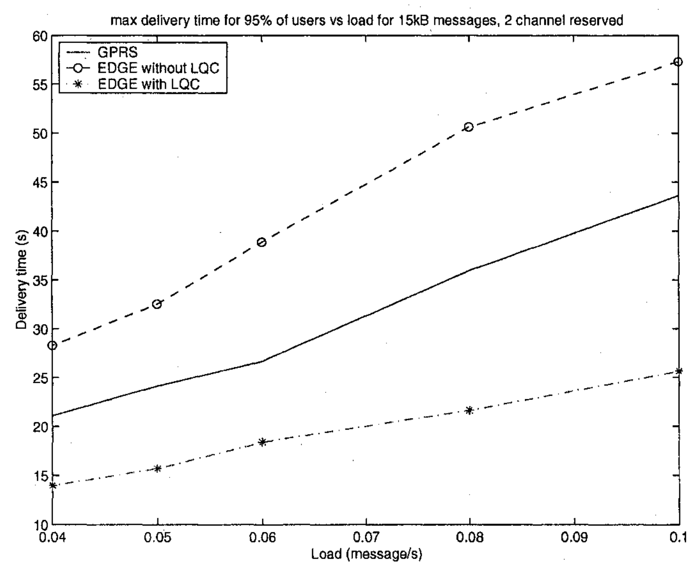


Figure A.15: Maximum delivery time for 95% of users, 15KB messages, 0.06 call/s, call length 60 s

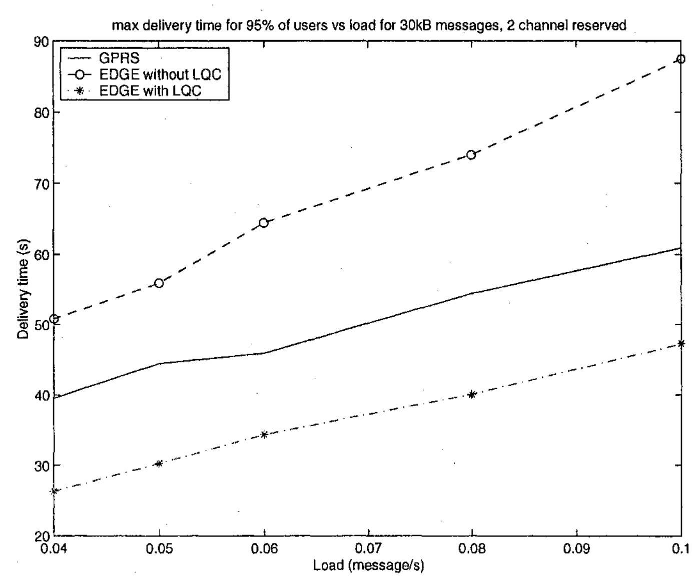


Figure A.16: Maximum delivery time for 95% of users, 30KB messages, 0.06 call/s, call length 60 s

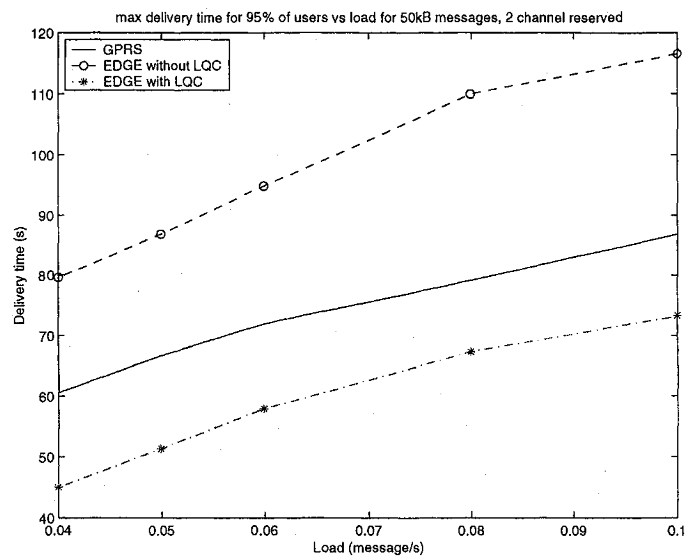


Figure A.17: Maximum delivery time for 95% of users, 50KB messages, 0.06 call/s, call length 60 s

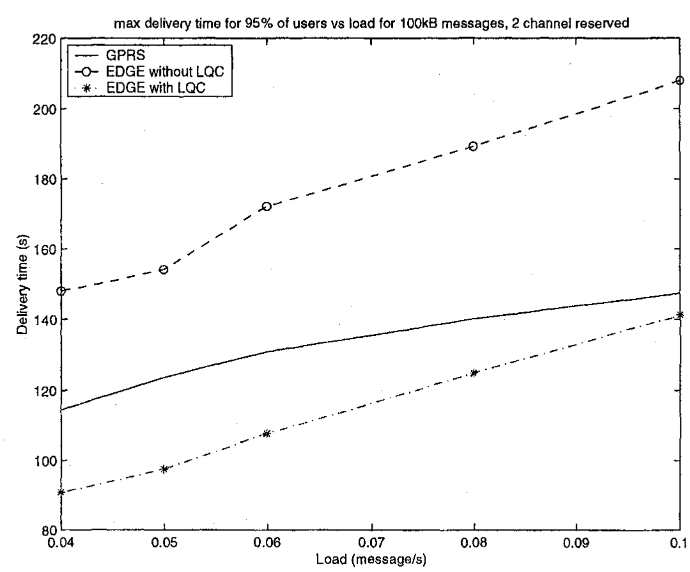


Figure A.18: Maximum delivery time for 95% of users, 100KB messages, 0.06 call/s, call length 60 s

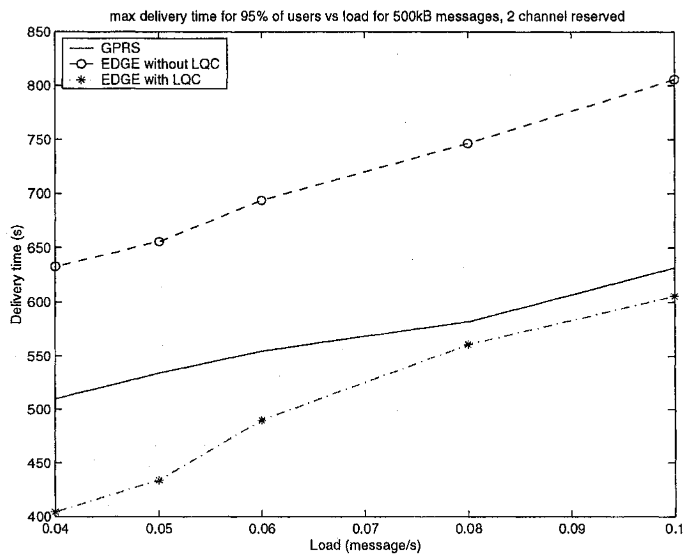


Figure A.19: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s

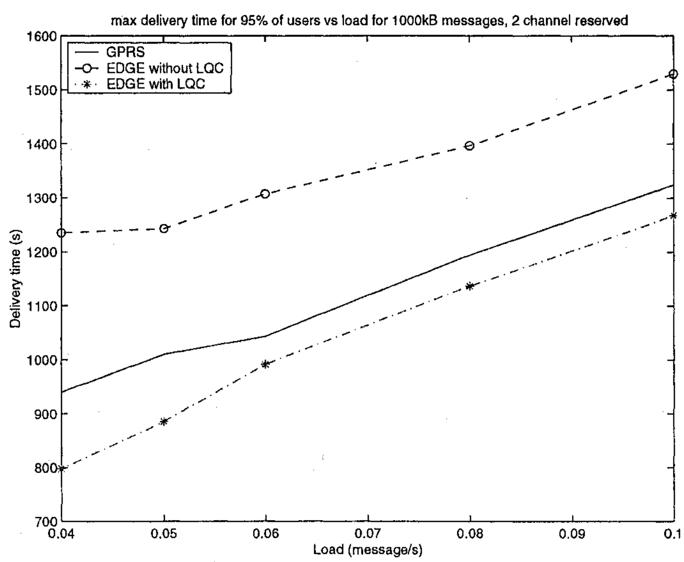


Figure A.20: Maximum delivery time for 95% of users, 1000KB messages, 0.06 call/s, call length 60 s

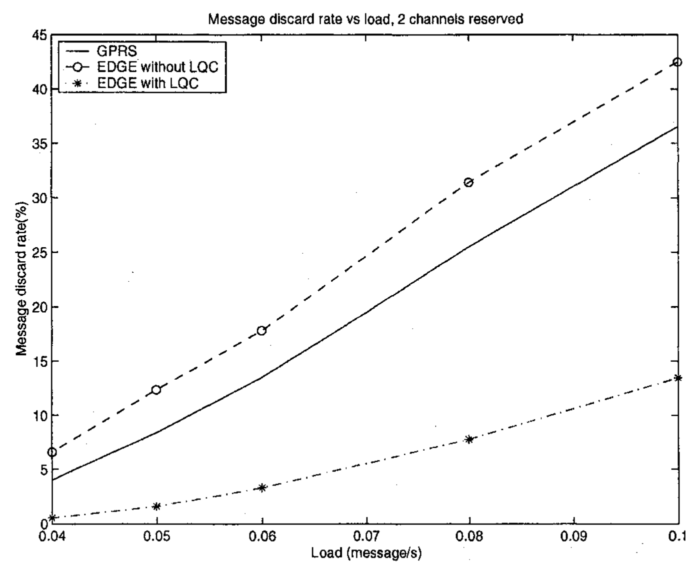


Figure A.21: Message discard rate, 0.06 call/s, call length 60 s

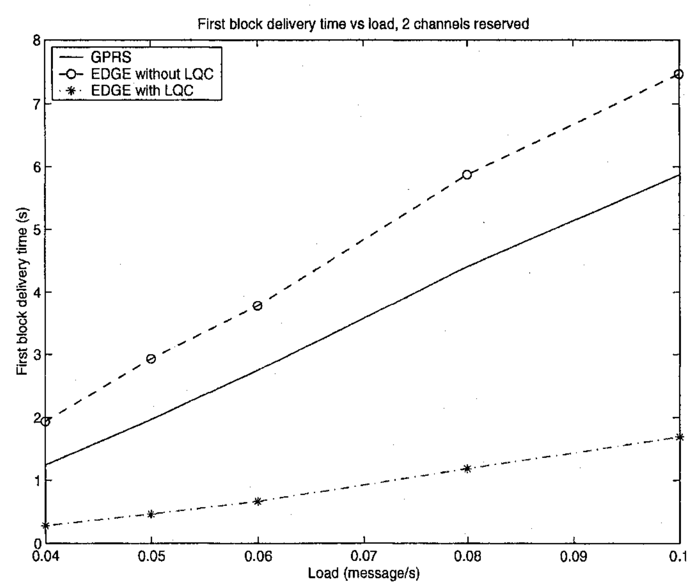


Figure A.22: First block delivery time, 0.06 call/s, call length 60 s

Appendix B

Results for different data channel reservation policies

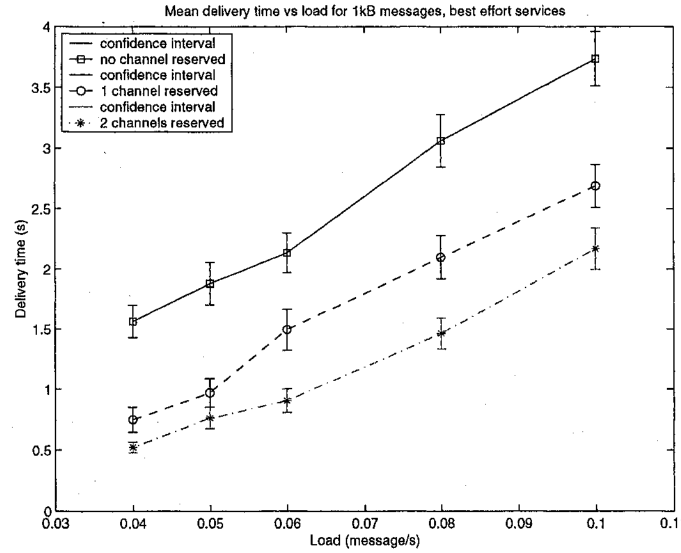


Figure B.1: Mean delivery time, 1KB messages, 0.06 call/s, call length 60 s

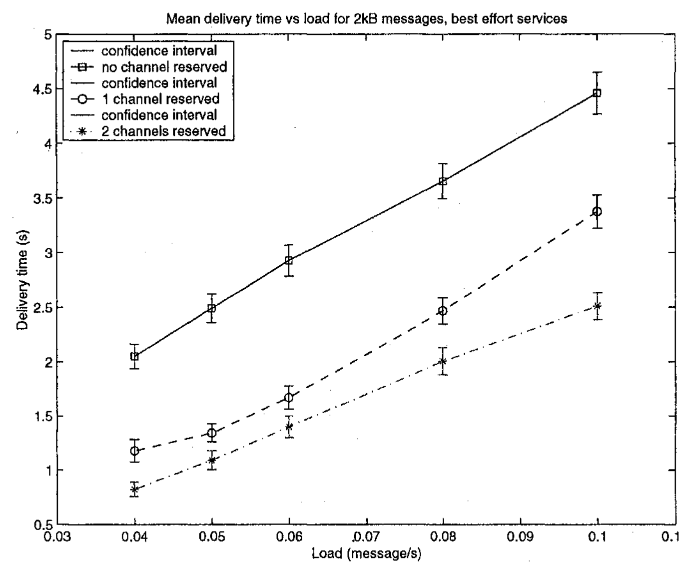


Figure B.2: Mean delivery time, 2KB messages, 0.06 call/s, call length 60 s

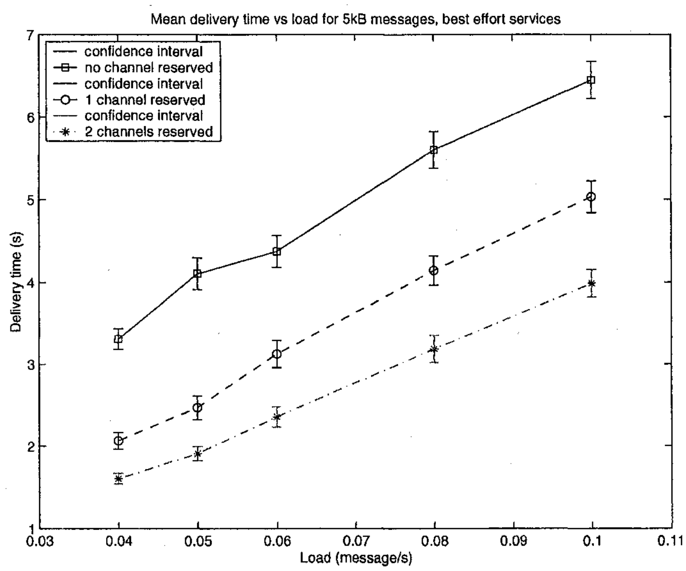


Figure B.3: Mean delivery time, 5KB messages, 0.06 call/s, call length 60 s

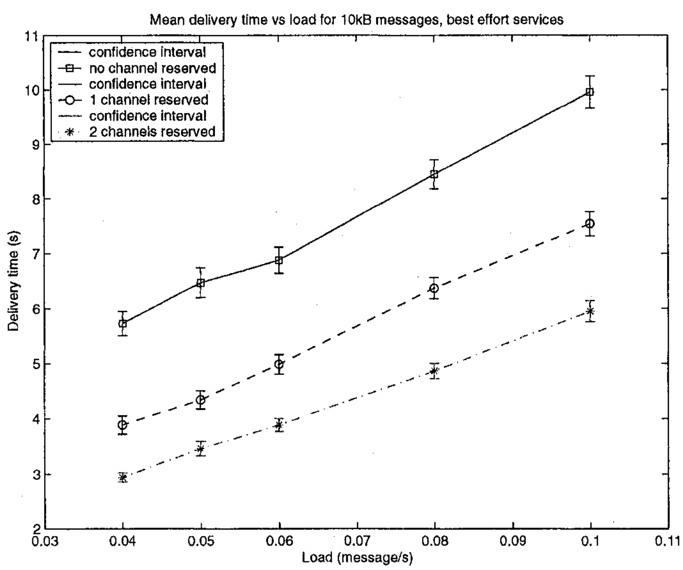


Figure B.4: Mean delivery time, 10KB messages, 0.06 call/s, call length 60 s

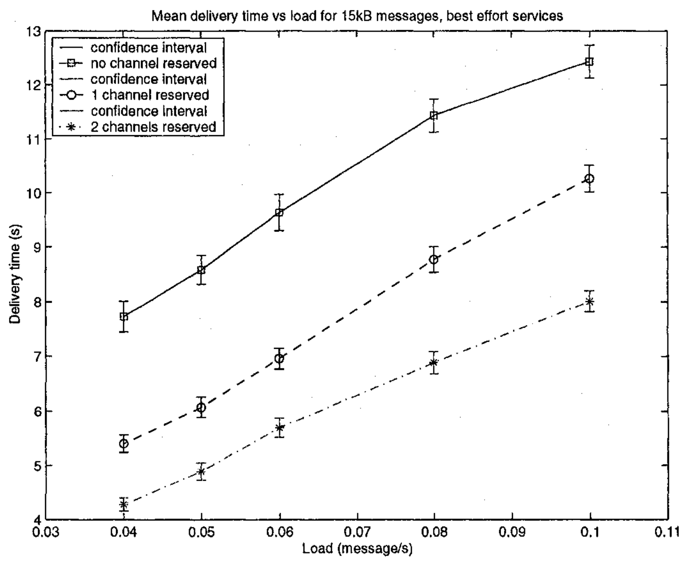


Figure B.5: Mean delivery time, 15KB messages, 0.06 call/s, call length 60 s

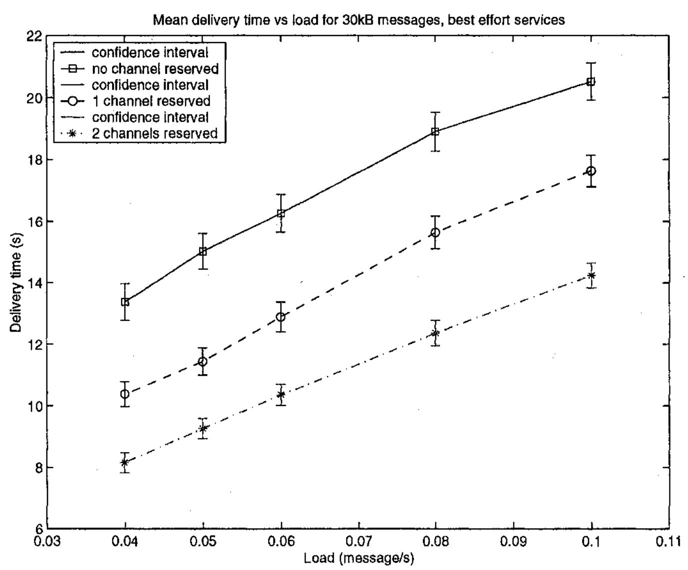


Figure B.6: Mean delivery time, 30KB messages, 0.06 call/s, call length 60 s

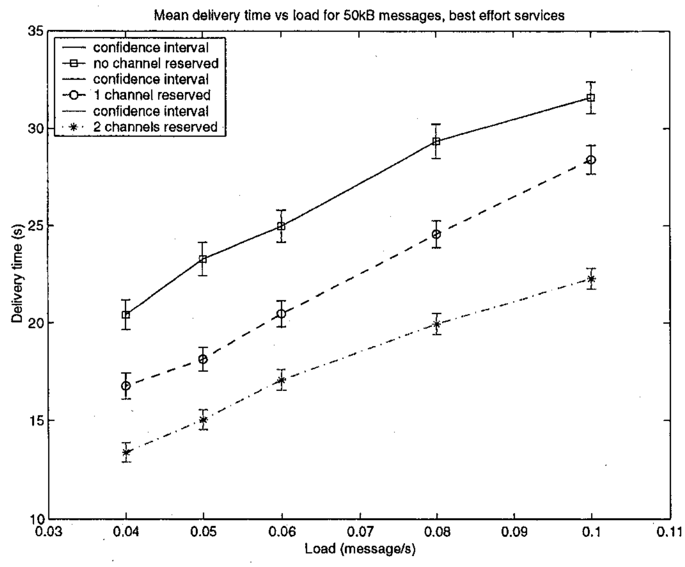


Figure B.7: Mean delivery time, 50KB messages, 0.06 call/s, call length 60 s

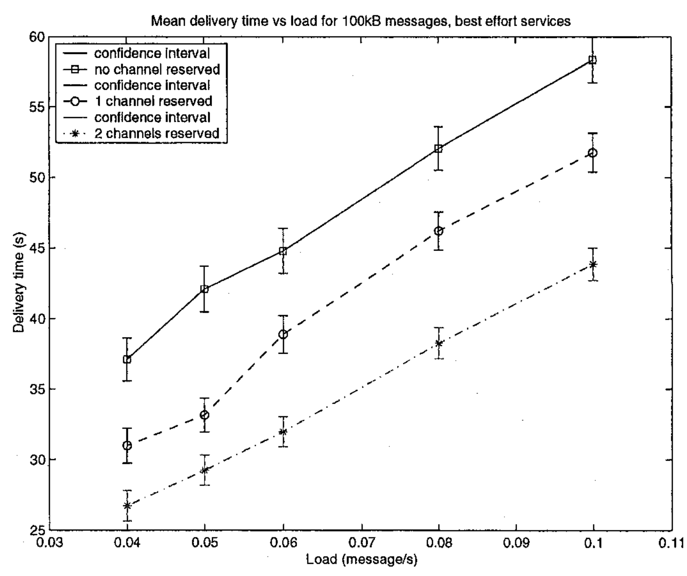


Figure B.8: Mean delivery time, 100KB messages, 0.06 call/s, call length 60 s

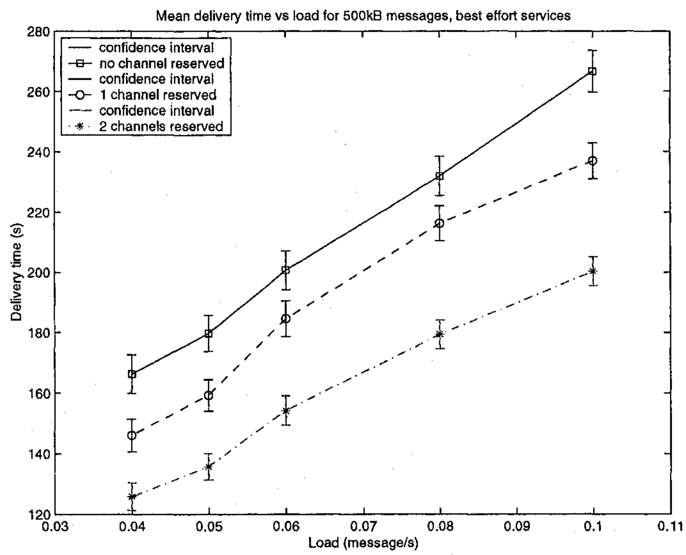


Figure B.9: Mean delivery time, 500KB messages, 0.06 call/s, call length 60 s

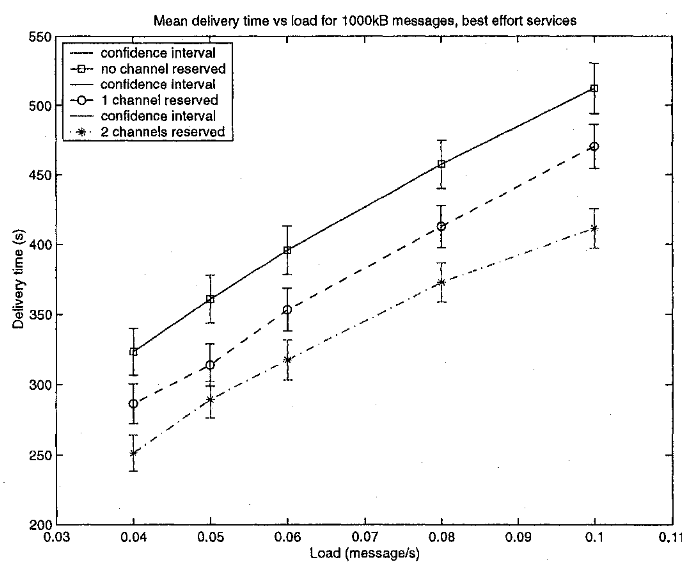


Figure B.10: Mean delivery time, 1000KB messages, 0.06 call/s, call length 60 s

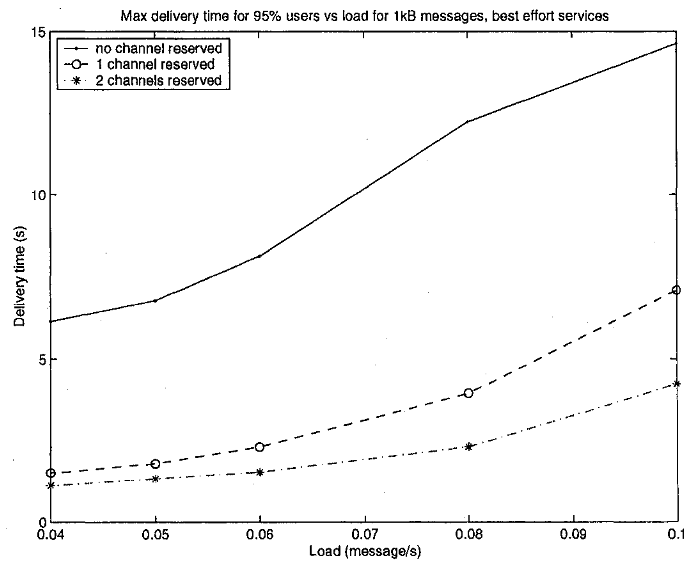


Figure B.11: Maximum delivery time for 95% of users, 1KB messages, 0.06 call/s, call length 60 s

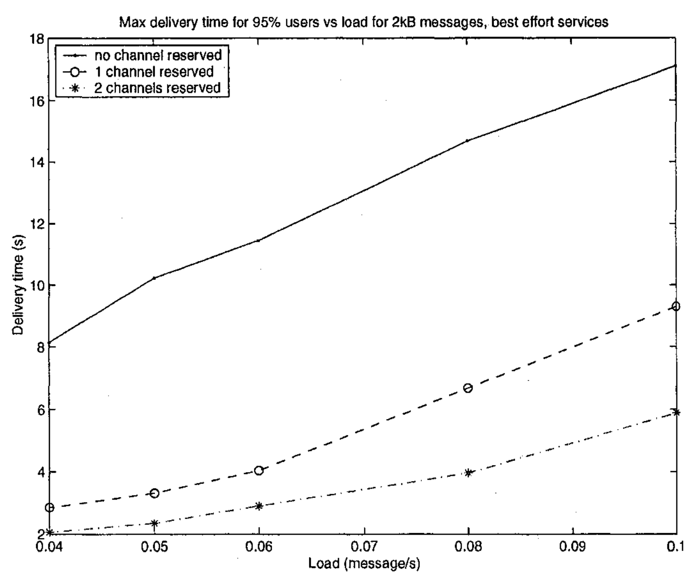


Figure B.12: Maximum delivery time for 95% of users, 2KB messages, 0.06 call/s, call length 60 s

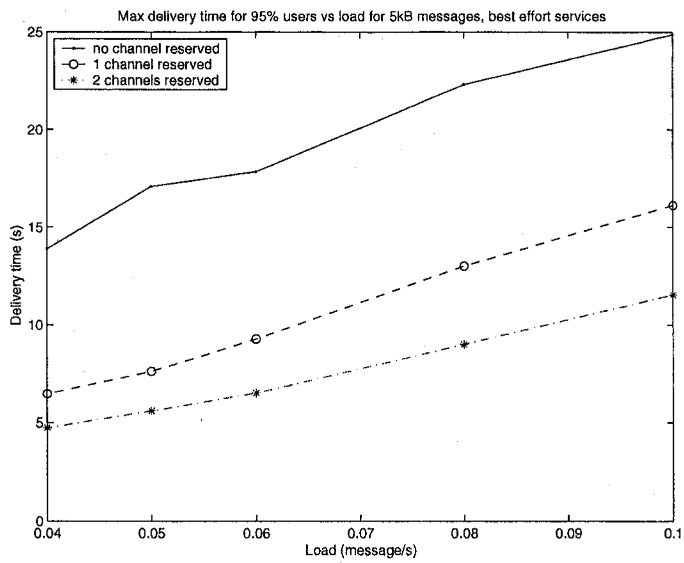


Figure B.13: Maximum delivery time for 95% of users, 5KB messages, 0.06 call/s, call length 60 s

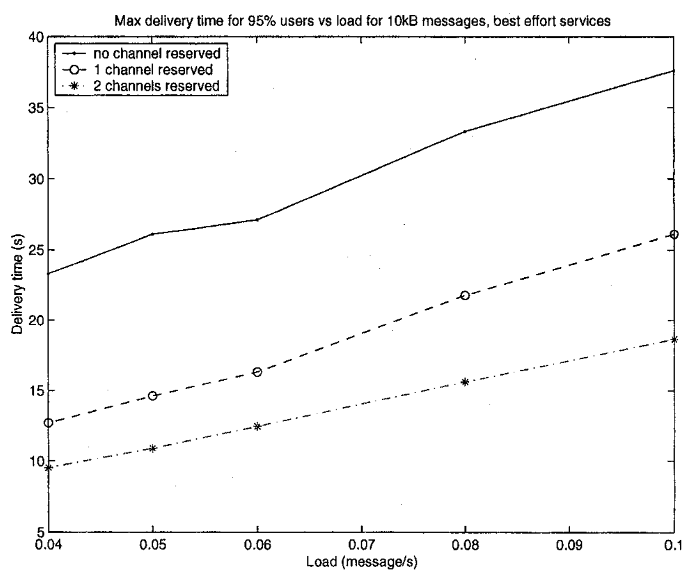


Figure B.14: Maximum delivery time for 95% of users, 10KB messages, 0.06 call/s, call length 60 s

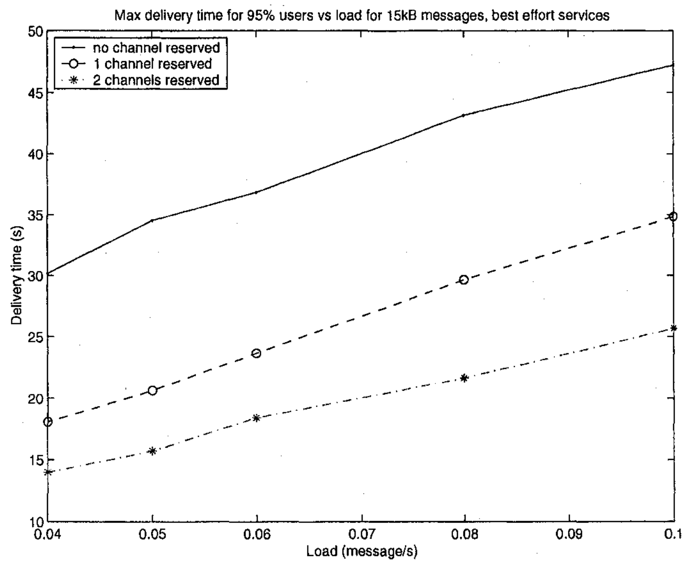


Figure B.15: Maximum delivery time for 95% of users, 15KB messages, 0.06 call/s, call length 60 s

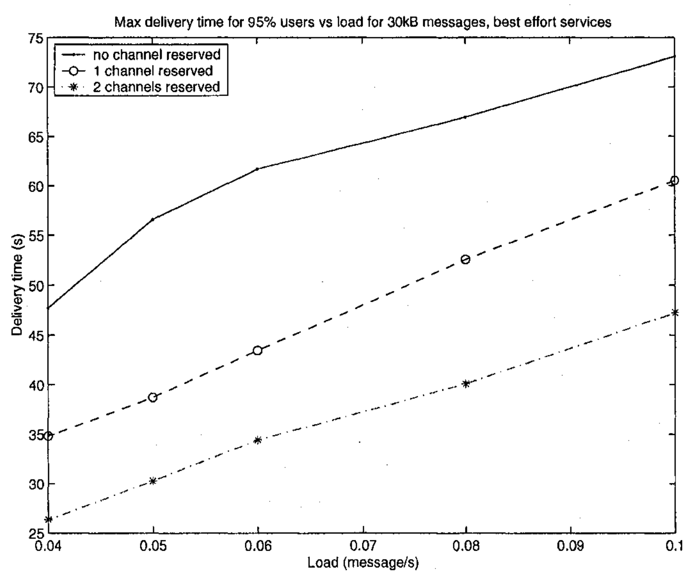


Figure B.16: Maximum delivery time for 95% of users, 30KB messages, 0.06 call/s, call length 60 s

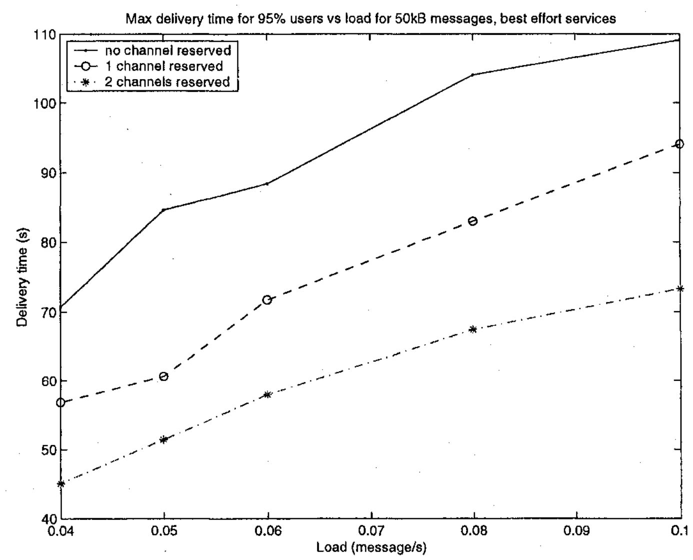


Figure B.17: Maximum delivery time for 95% of users, 50KB messages, 0.06 call/s, call length 60 s

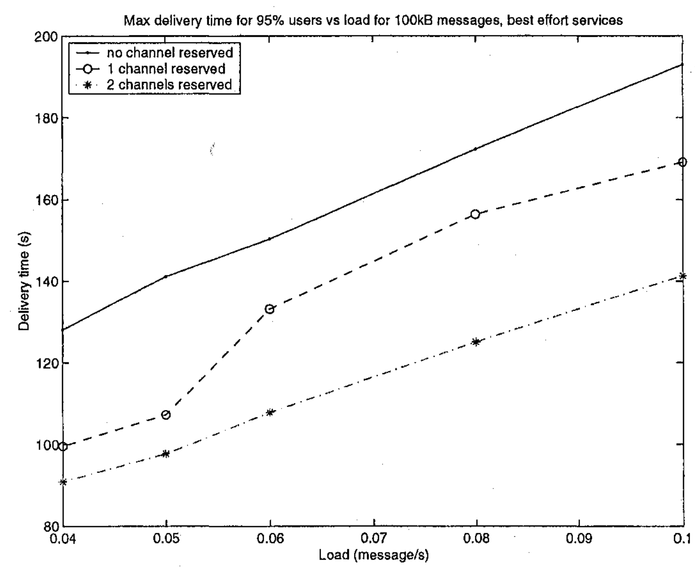


Figure B.18: Maximum delivery time for 95% of users, 100KB messages, 0.06 call/s, call length 60 s

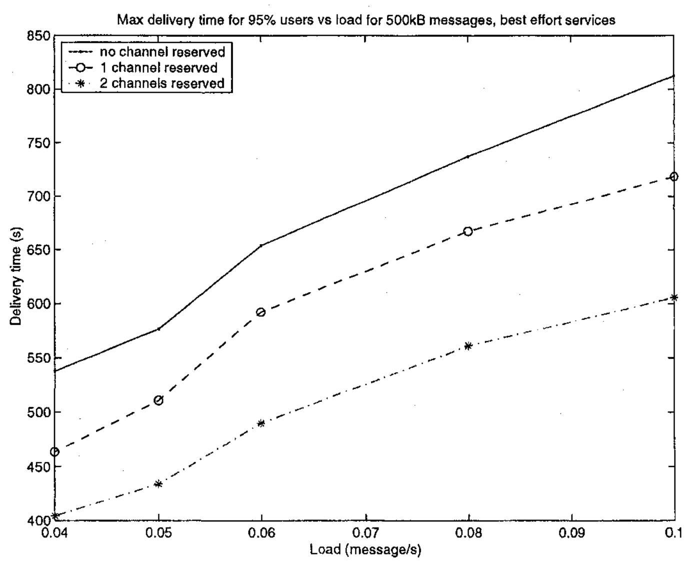


Figure B.19: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s

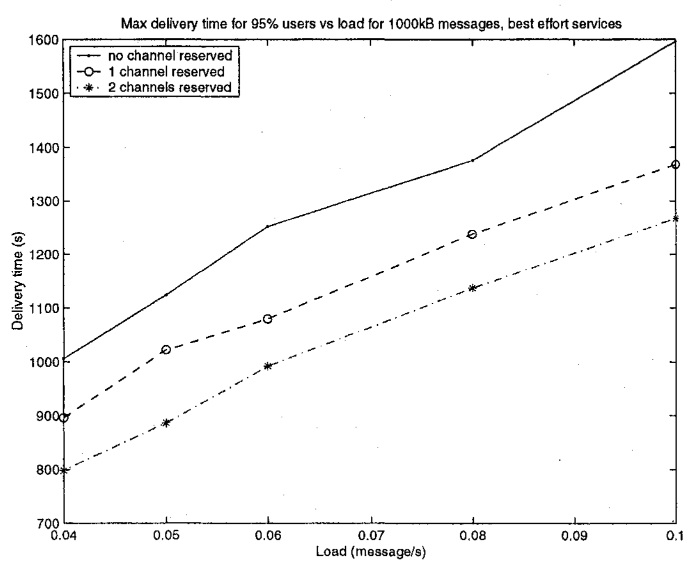


Figure B.20: Maximum delivery time for 95% of users, 1000KB messages, 0.06 call/s, call length 60 s

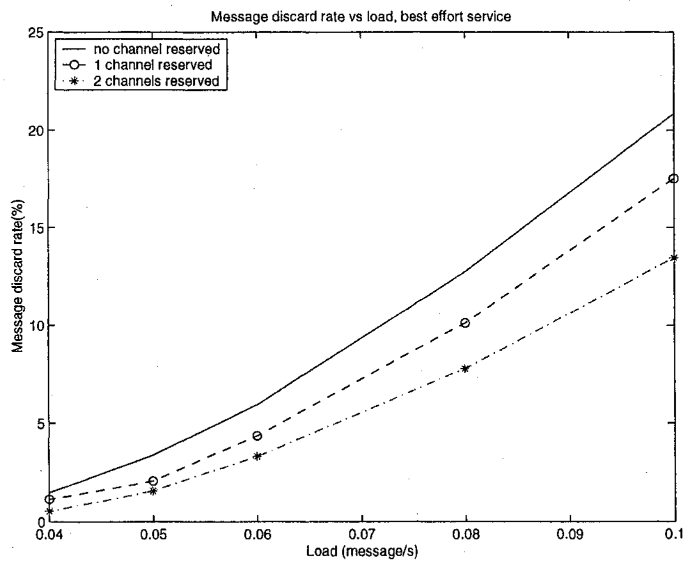


Figure B.21: Message discard rate, 0.06 call/s, call length 60 s

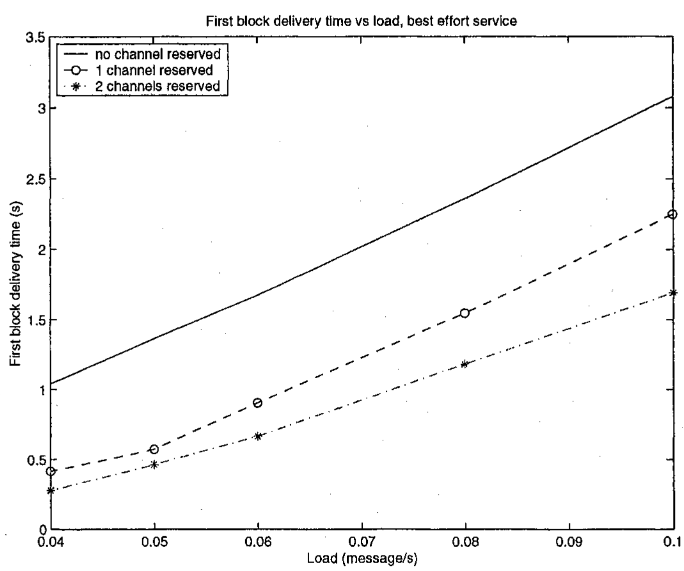


Figure B.22: First block delivery time , 0.06 call/s, call length 60 s

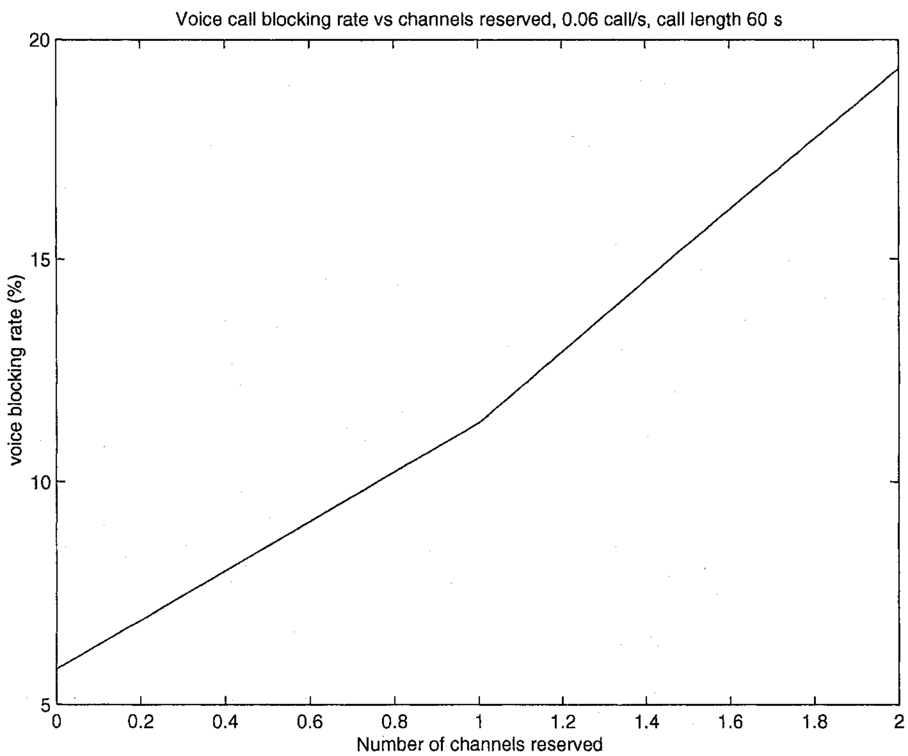


Figure B.23: Voice blocking rate , 0.06 call/s, call length 60 s

Appendix C

Results for different LQC Algorithms

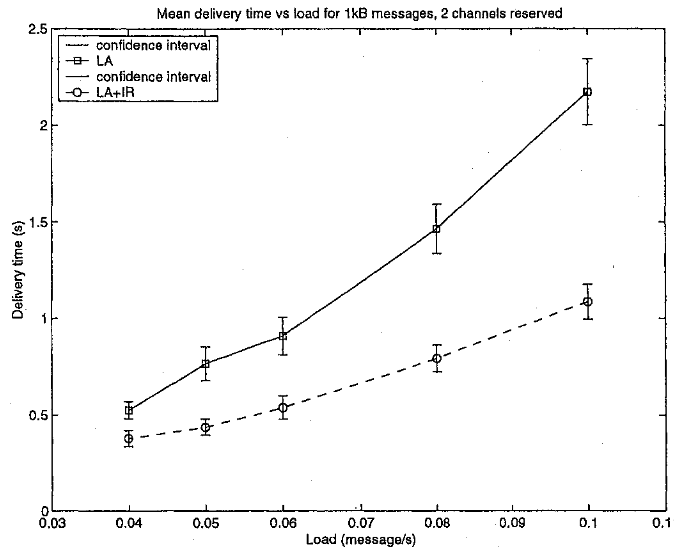


Figure C.1: Mean delivery time, 1KB messages, 0.06 call/s, call length 60 s

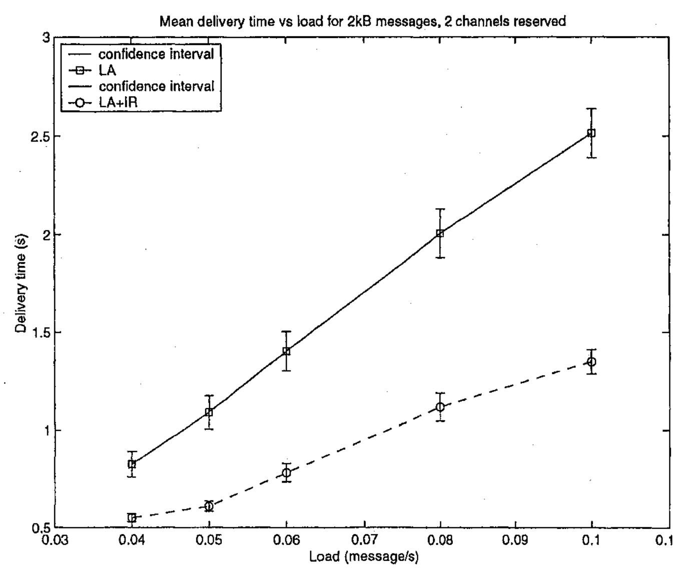


Figure C.2: Mean delivery time, 2KB messages, 0.06 call/s, call length 60 s

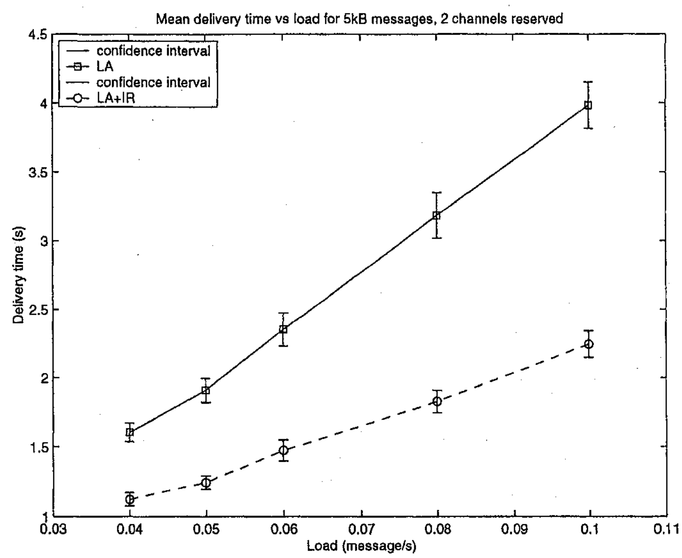


Figure C.3: Mean delivery time, 5KB messages, 0.06 call/s, call length 60 s

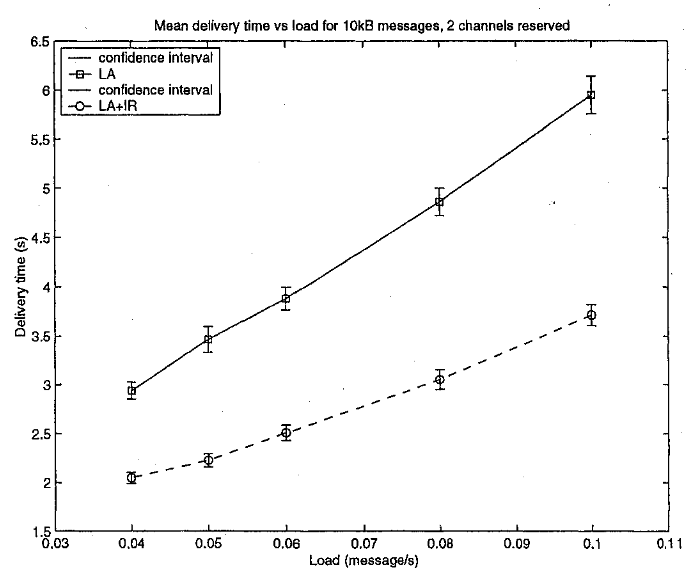


Figure C.4: Mean delivery time, 10KB messages, 0.06 call/s, call length 60 s

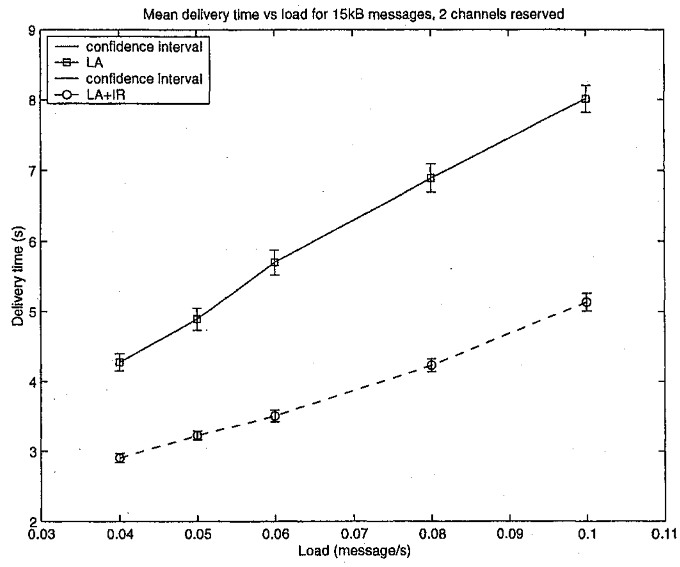


Figure C.5: Mean delivery time, 15KB messages, 0.06 call/s, call length 60 s

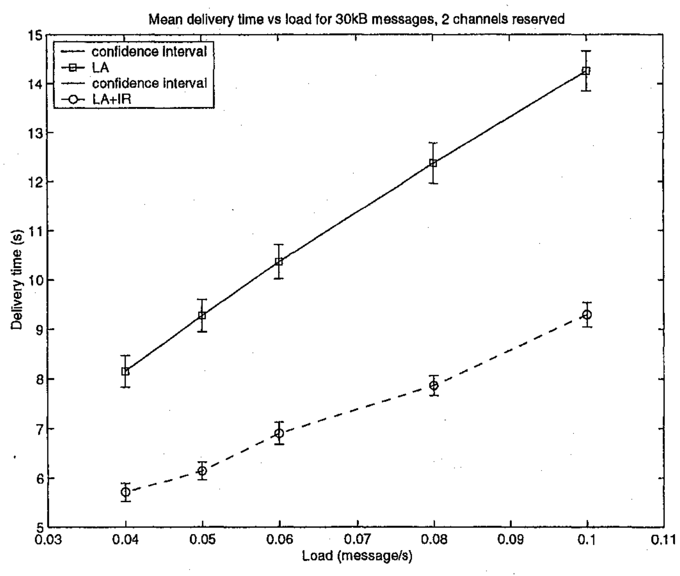


Figure C.6: Mean delivery time, 30KB messages, 0.06 call/s, call length 60 s

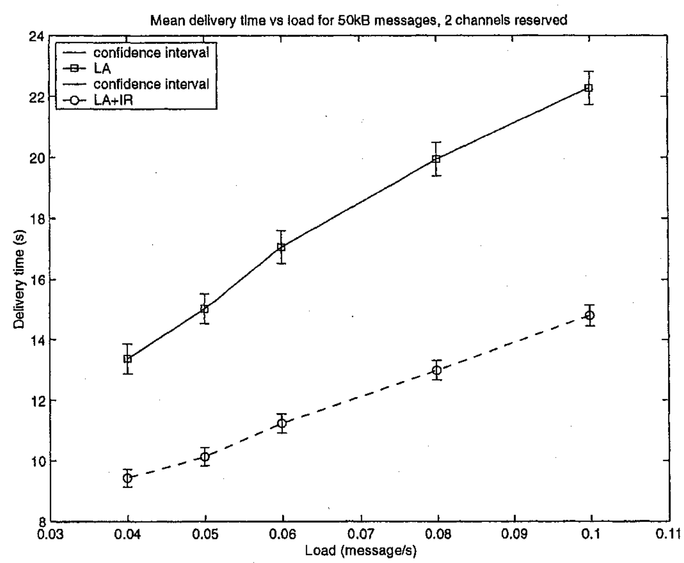


Figure C.7: Mean delivery time, 50KB messages, 0.06 call/s, call length 60 s

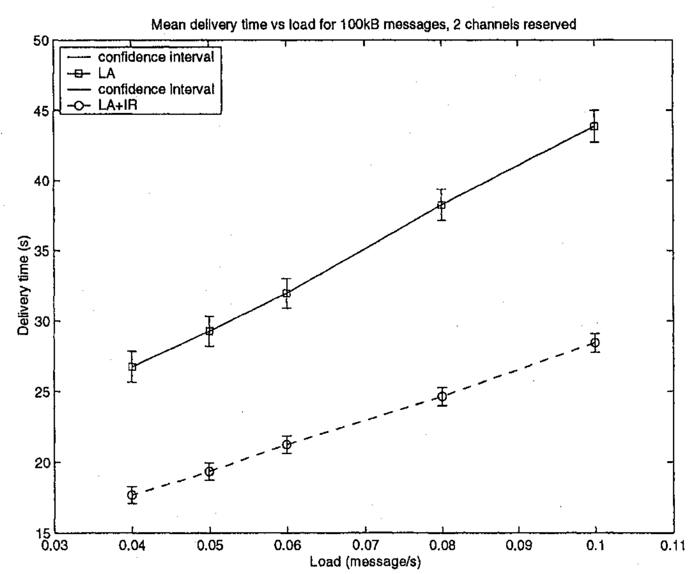


Figure C.8: Mean delivery time, 100KB messages, 0.06 call/s, call length 60 s

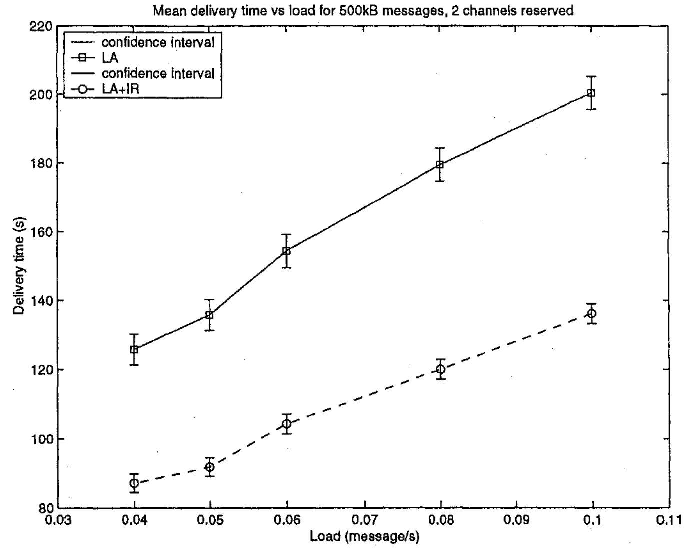


Figure C.9: Mean delivery time, 500KB messages, 0.06 call/s, call length 60 s

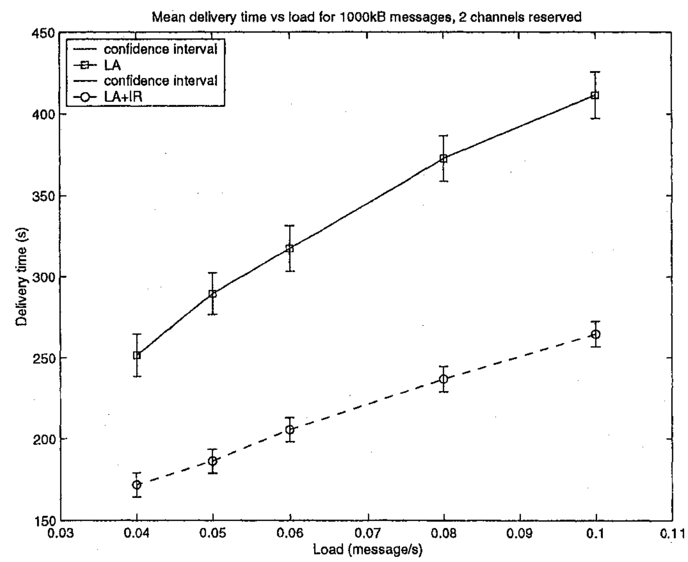


Figure C.10: Mean delivery time, 1000KB messages, 0.06 call/s, call length 60 s

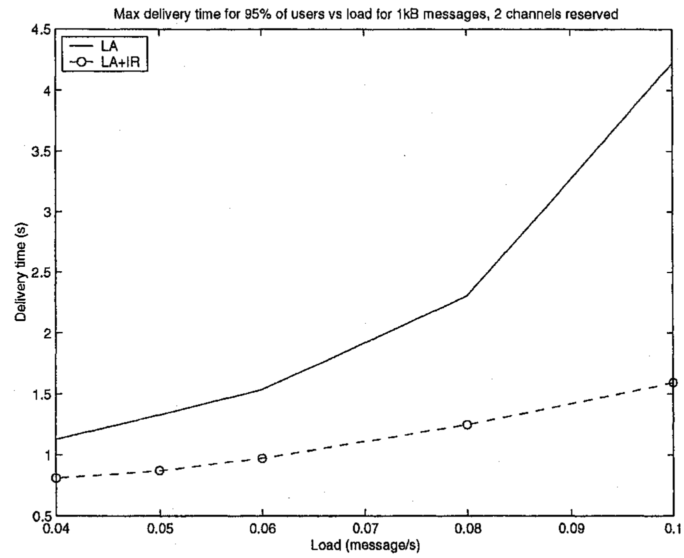


Figure C.11: Maximum delivery time for 95% of users, 1KB messages, 0.06 call/s, call length 60 s

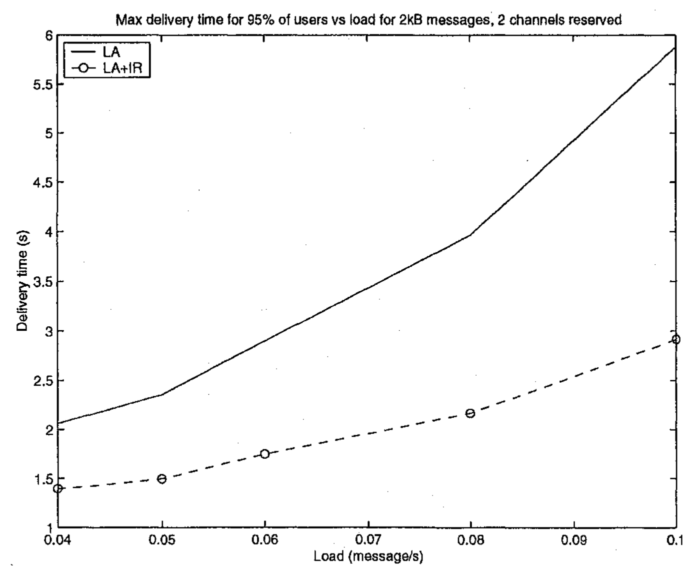


Figure C.12: Maximum delivery time for 95% of users, 2KB messages, 0.06 call/s, call length 60 s

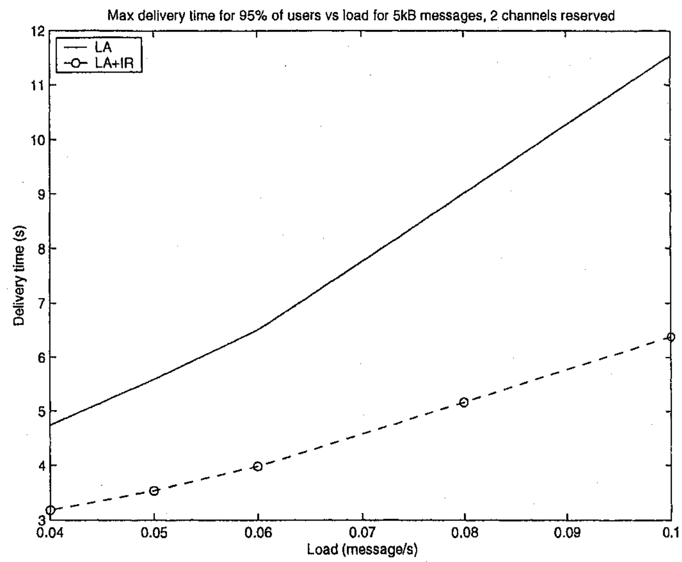


Figure C.13: Maximum delivery time for 95% of users, 5KB messages, 0.06 call/s, call length 60 s

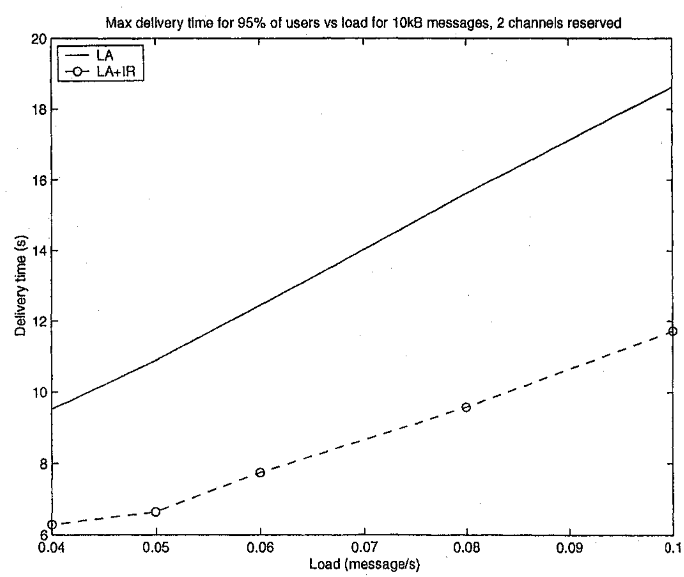


Figure C.14: Maximum delivery time for 95% of users, 10KB messages, 0.06 call/s, call length 60 s

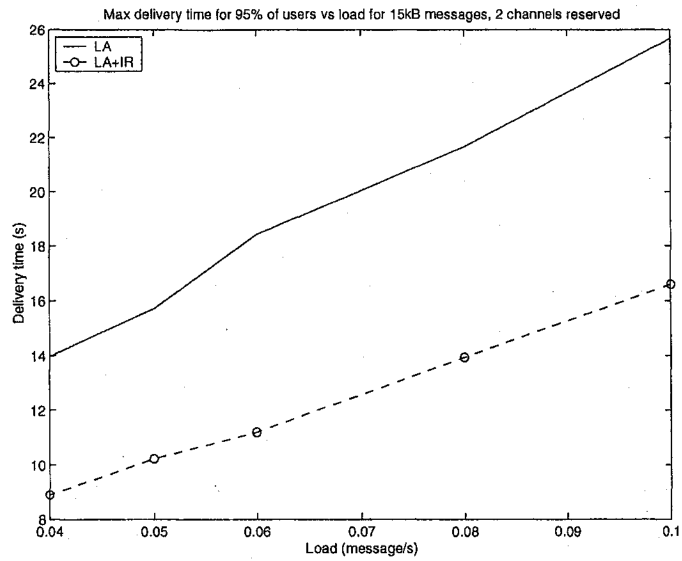


Figure C.15: Maximum delivery time for 95% of users, 15KB messages, 0.06 call/s, call length 60 s

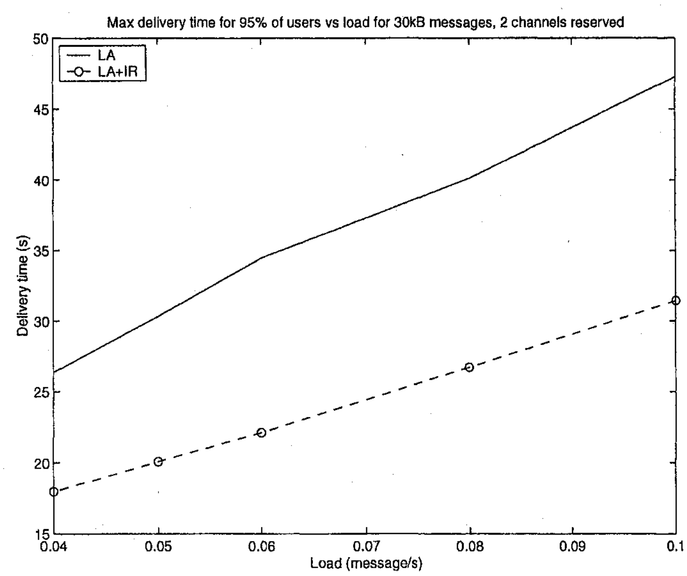


Figure C.16: Maximum delivery time for 95% of users, 30KB messages, 0.06 call/s, call length 60 s

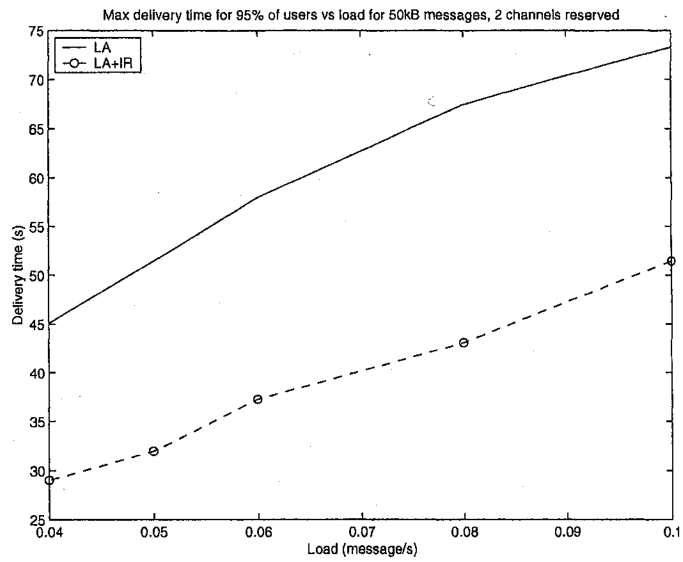


Figure C.17: Maximum delivery time for 95% of users, 50KB messages, 0.06 call/s, call length 60 s

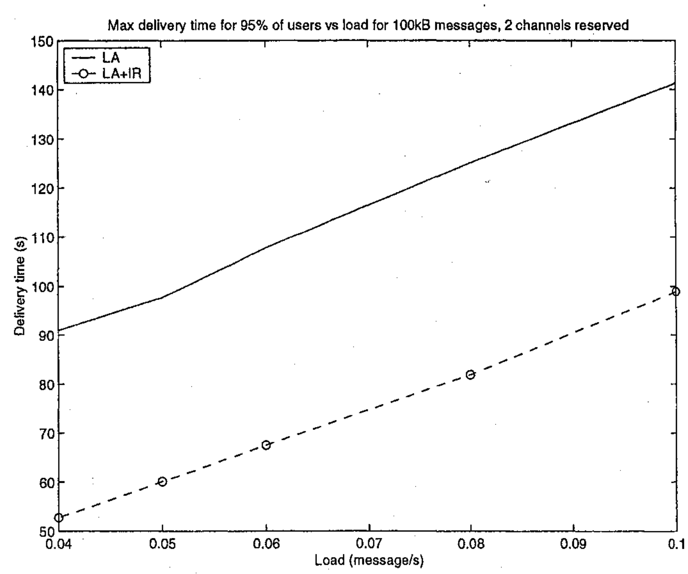


Figure C.18: Maximum delivery time for 95% of users, 100KB messages, 0.06 call/s, call length 60 s

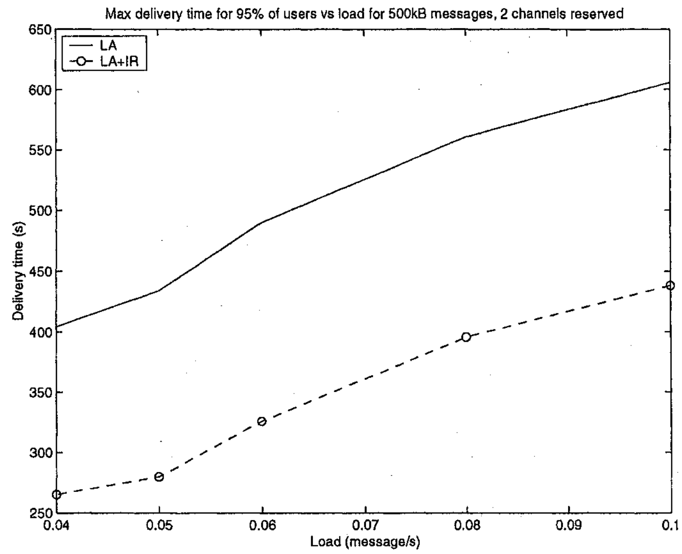


Figure C.19: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s

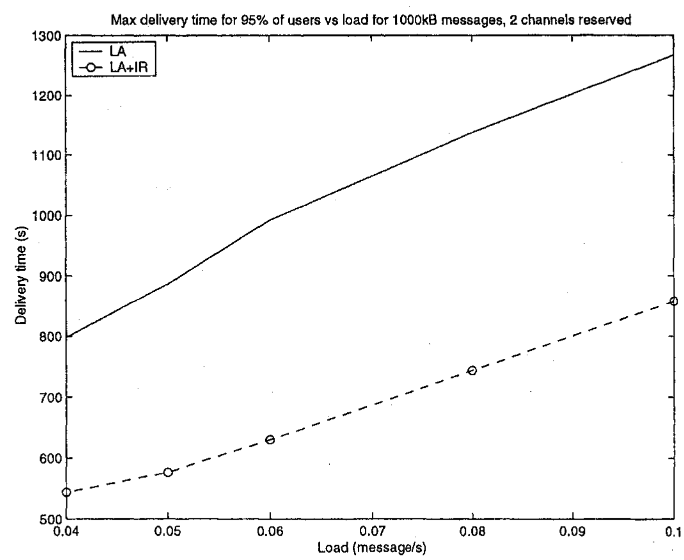


Figure C.20: Maximum delivery time for 95% of users, 1000KB messages, 0.06 call/s, call length 60 s

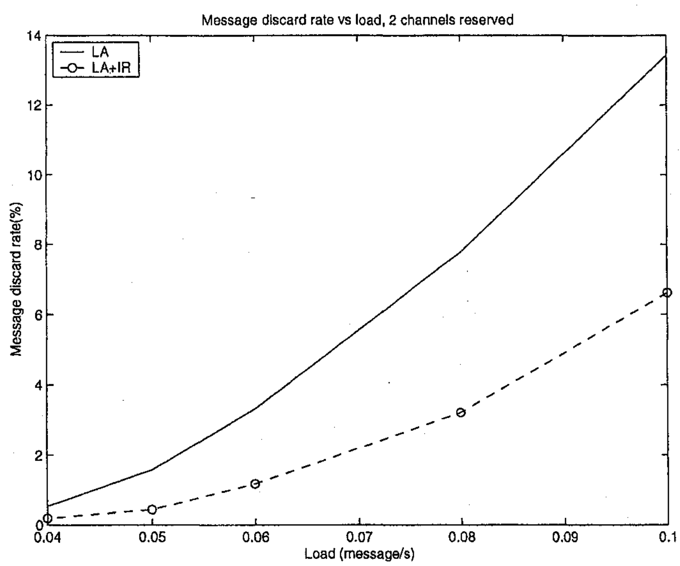


Figure C.21: Message discard rate, 0.06 call/s, call length 60 s

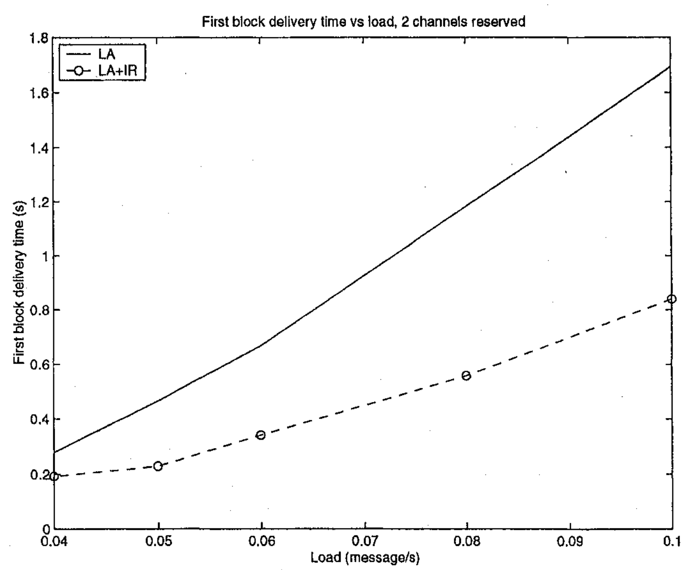


Figure C.22: First block delivery time , 0.06 call/s, call length 60 s

Appendix D

Results on the effects of admission control (AC)

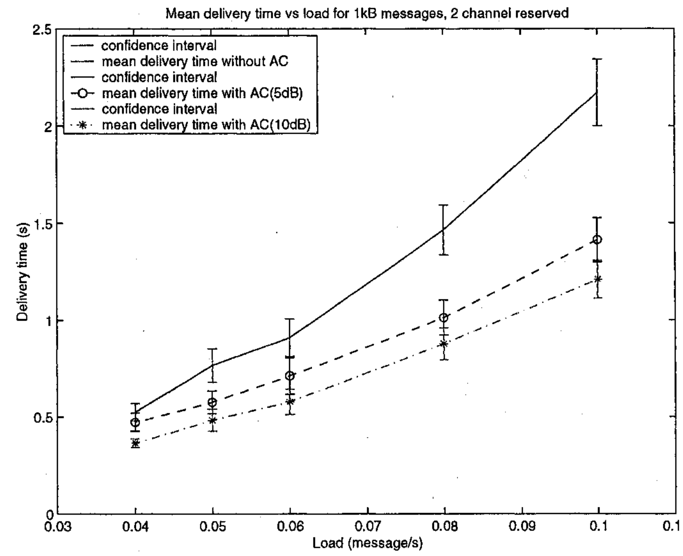


Figure D.1: Mean delivery time, 1KB messages, 0.06 call/s, call length 60 s

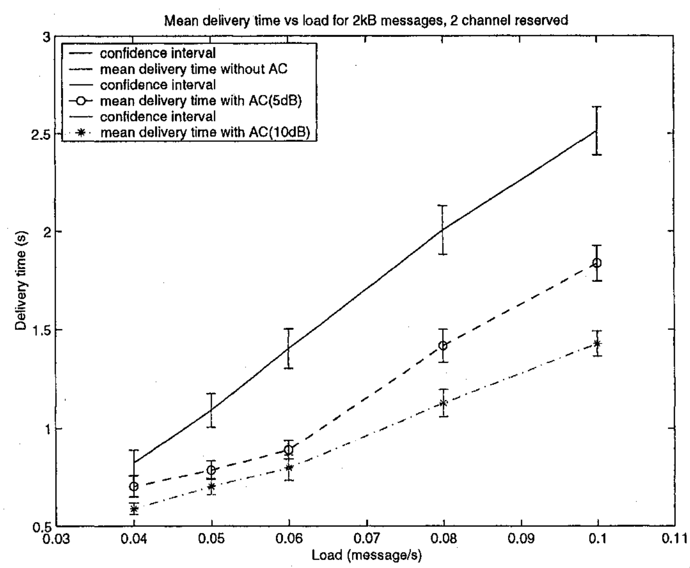


Figure D.2: Mean delivery time, 2KB messages, 0.06 call/s, call length 60 s

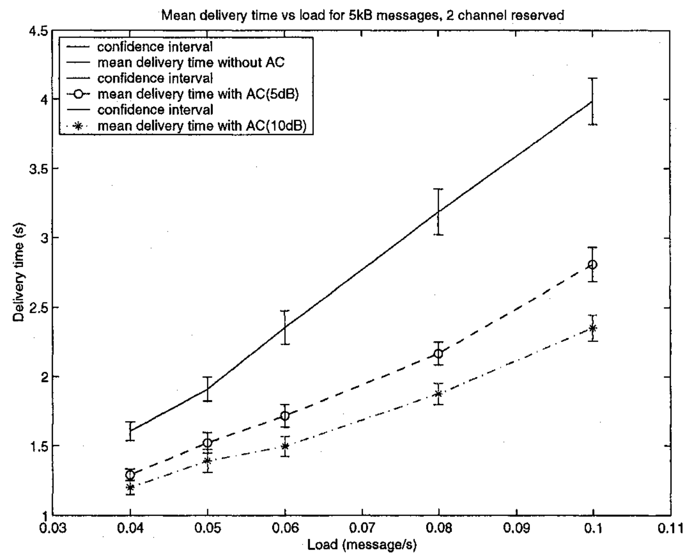


Figure D.3: Mean delivery time, 5KB messages, 0.06 call/s, call length 60 s

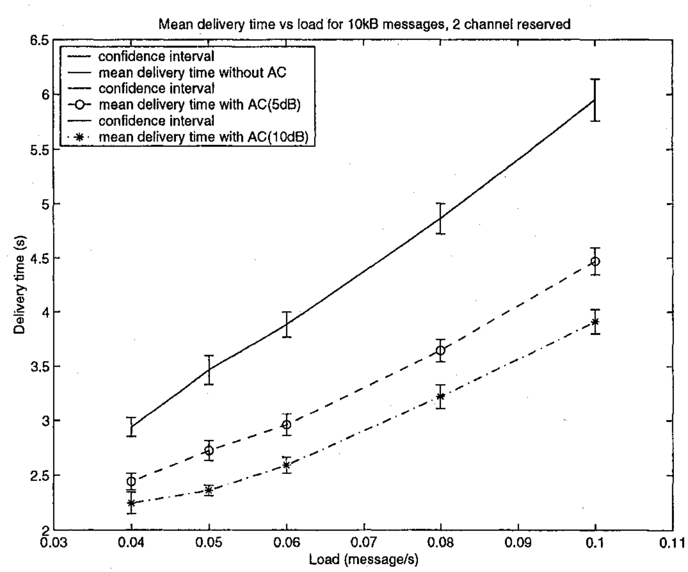


Figure D.4: Mean delivery time, 10KB messages, 0.06 call/s, call length 60 s

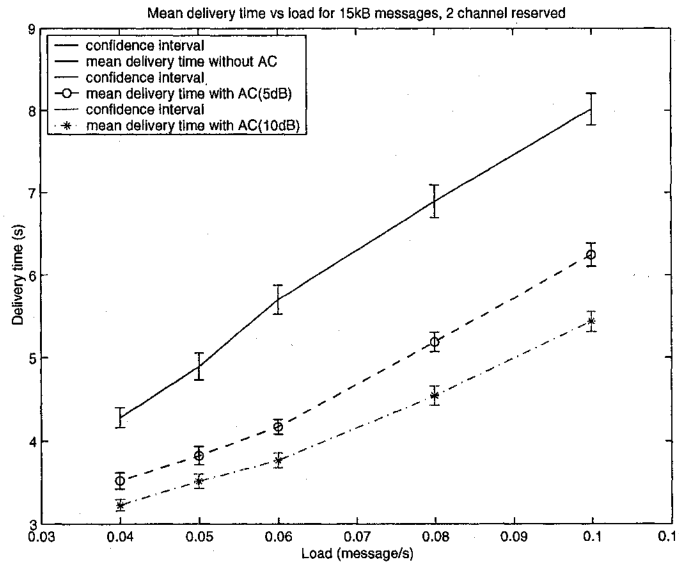


Figure D.5: Mean delivery time, 15KB messages, 0.06 call/s, call length 60 s

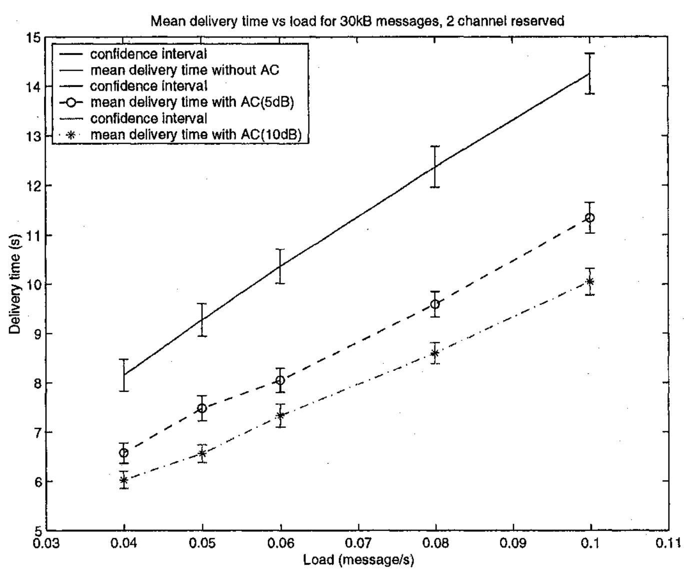


Figure D.6: Mean delivery time, 30KB messages, 0.06 call/s, call length 60 s

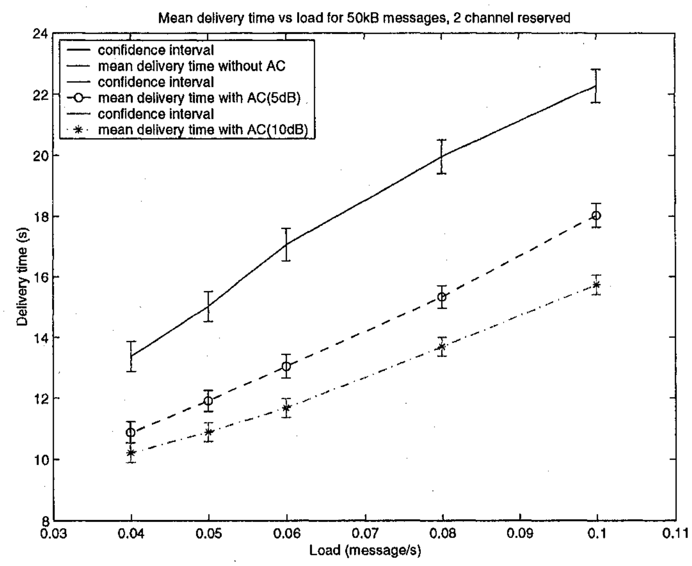


Figure D.7: Mean delivery time, 50KB messages, 0.06 call/s, call length 60 s

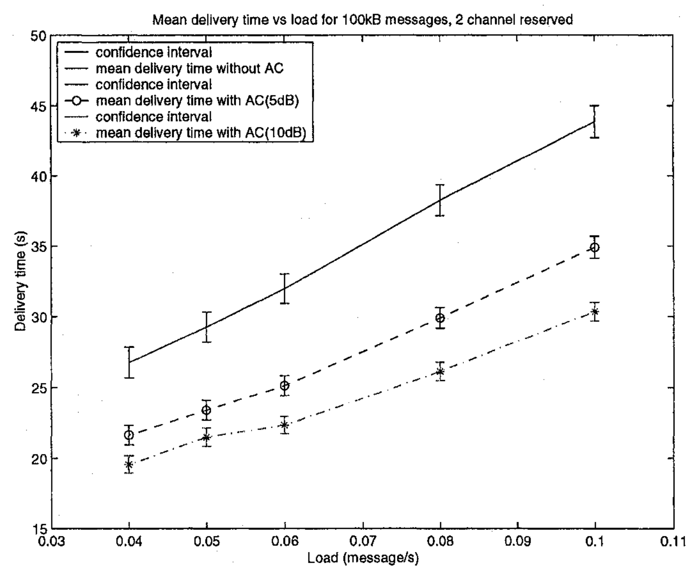


Figure D.8: Mean delivery time, 100KB messages, 0.06 call/s, call length 60 s

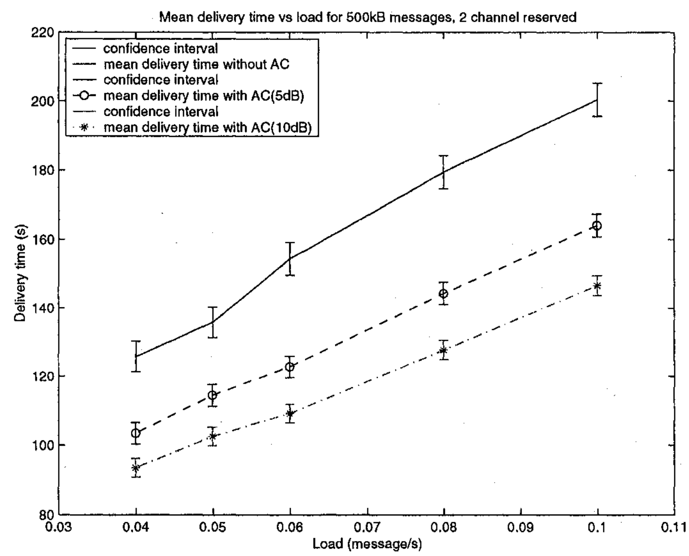


Figure D.9: Mean delivery time, 500KB messages, 0.06 call/s, call length 60 s

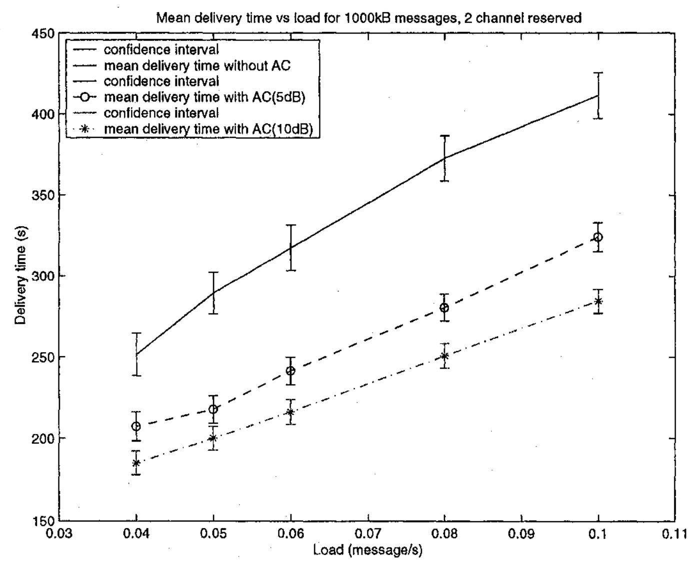


Figure D.10: Mean delivery time, 1000KB messages, 0.06 call/s, call length 60 s

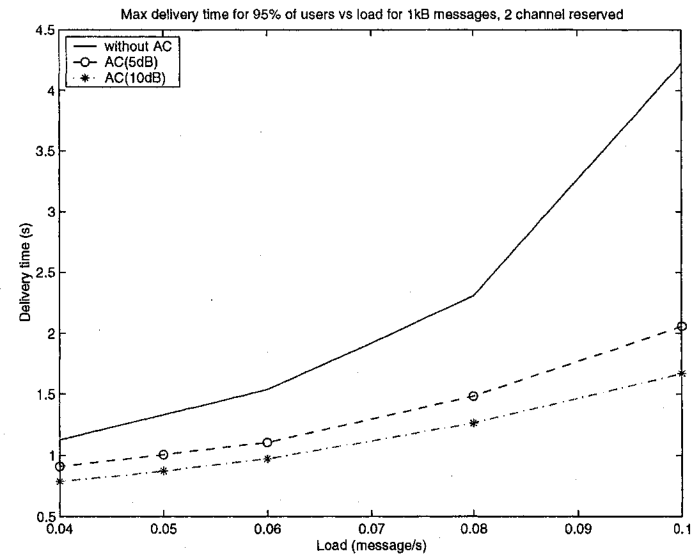


Figure D.11: Maximum delivery time for 95% of users, 1KB messages, 0.06 call/s, call length 60 s

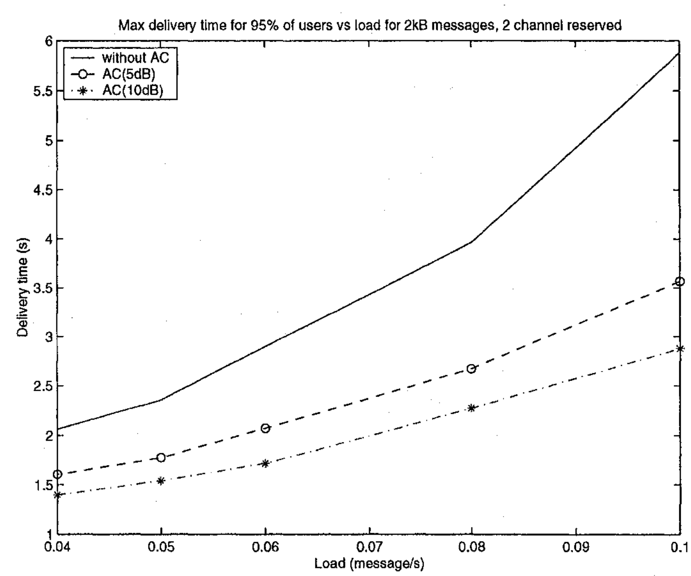


Figure D.12: Maximum delivery time for 95% of users, 2KB messages, 0.06 call/s, call length 60 s

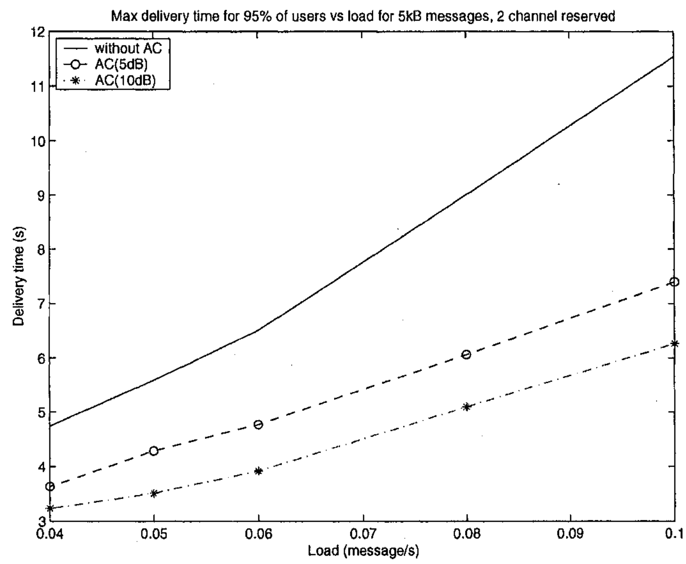


Figure D.13: Maximum delivery time for 95% of users, 5KB messages, 0.06 call/s, call length 60 s

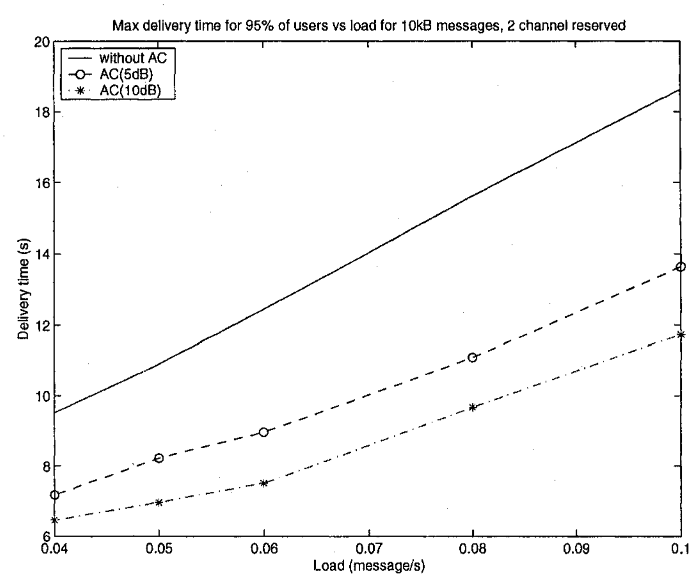


Figure D.14: Maximum delivery time for 95% of users, 10KB messages, 0.06 call/s, call length 60 s

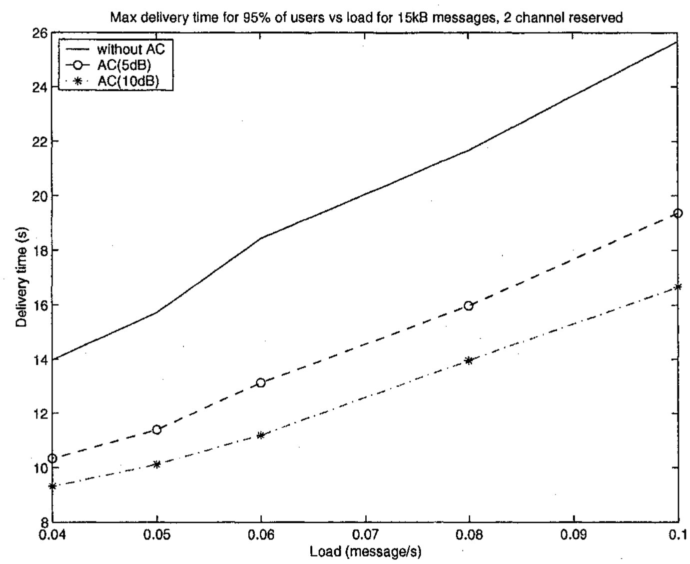


Figure D.15: Maximum delivery time for 95% of users, 15KB messages, 0.06 call/s, call length 60 s

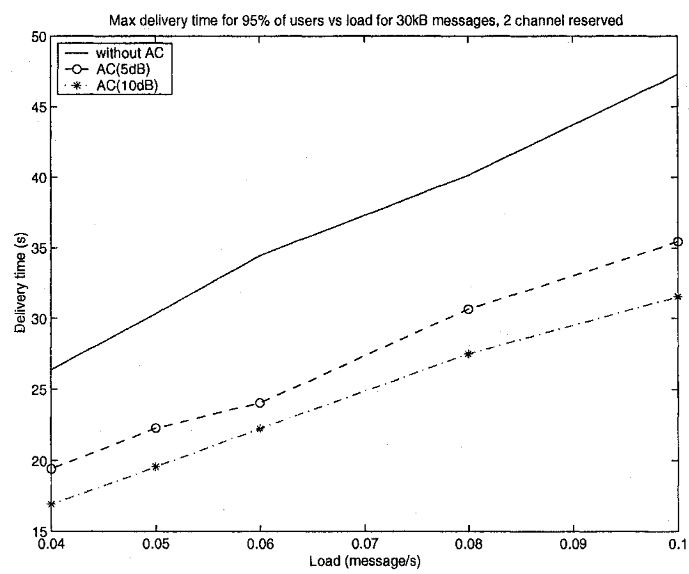


Figure D.16: Maximum delivery time for 95% of users, 30KB messages, 0.06 call/s, call length 60 s

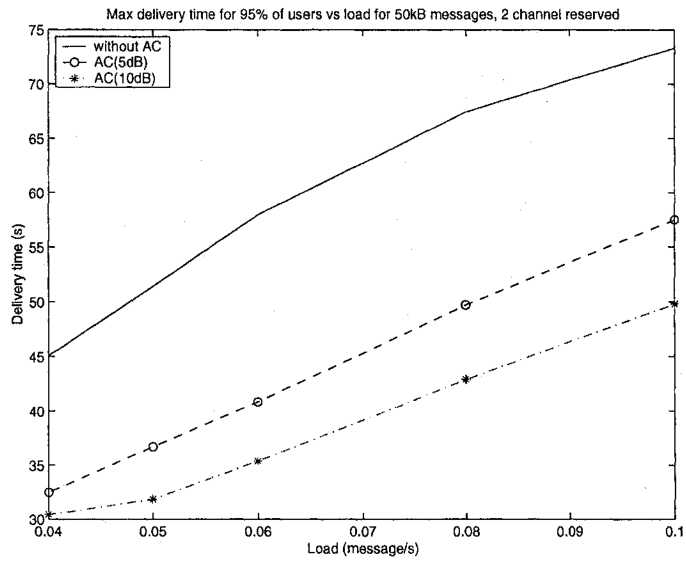


Figure D.17: Maximum delivery time for 95% of users, 50KB messages, 0.06 call/s, call length 60 s

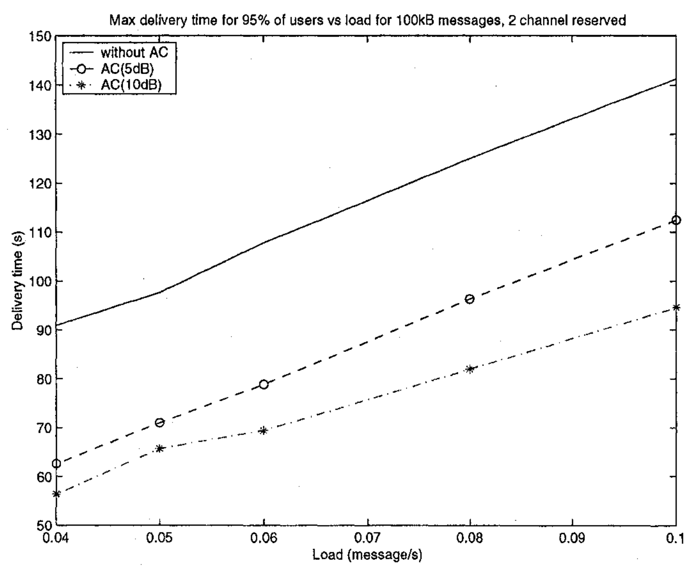


Figure D.18: Maximum delivery time for 95% of users, 100KB messages, 0.06 call/s, call length 60 s

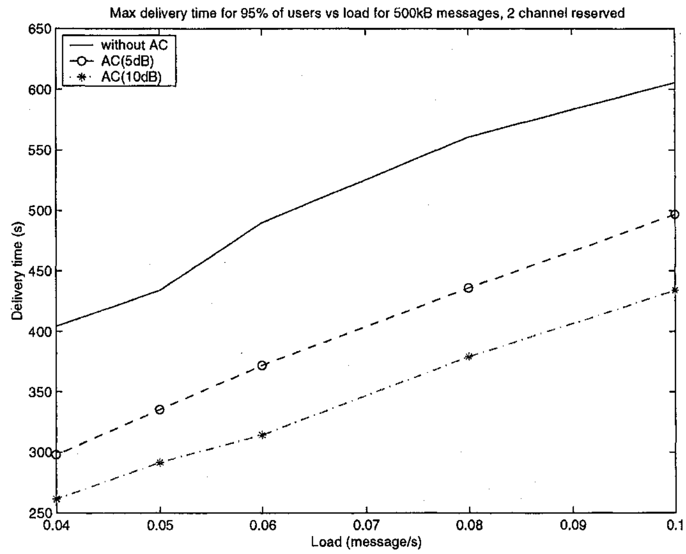


Figure D.19: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s

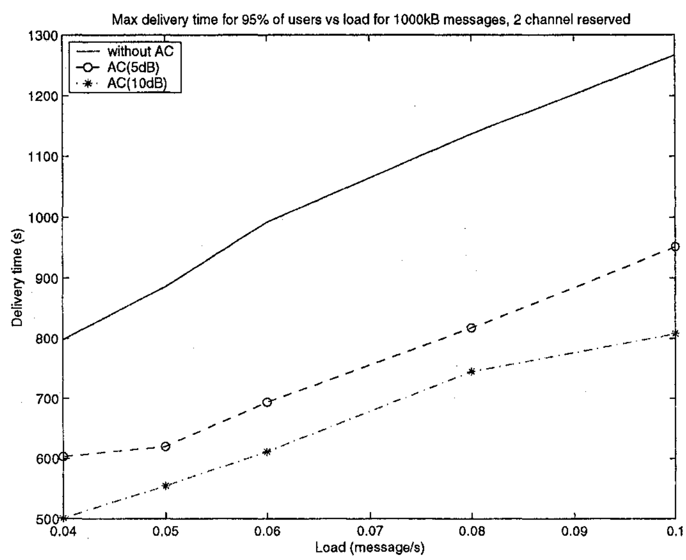


Figure D.20: Maximum delivery time for 95% of users, 1000KB messages, 0.06 call/s, call length 60 s

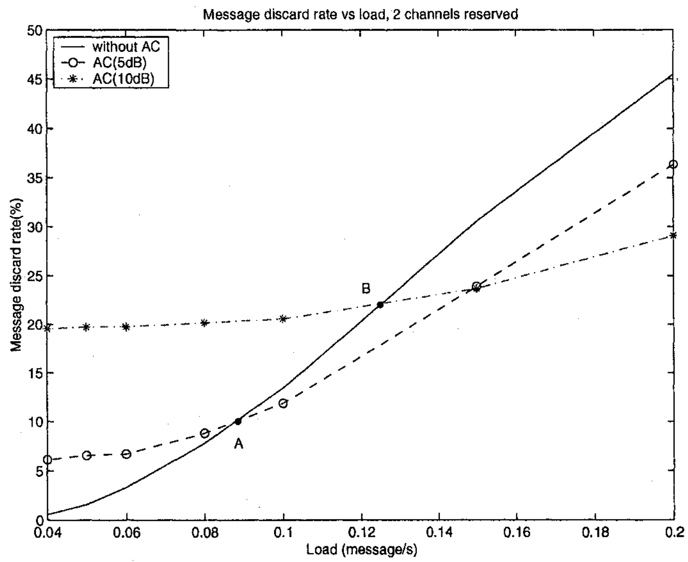


Figure D.21: Message discard rate, 0.06 call/s, call length 60 s

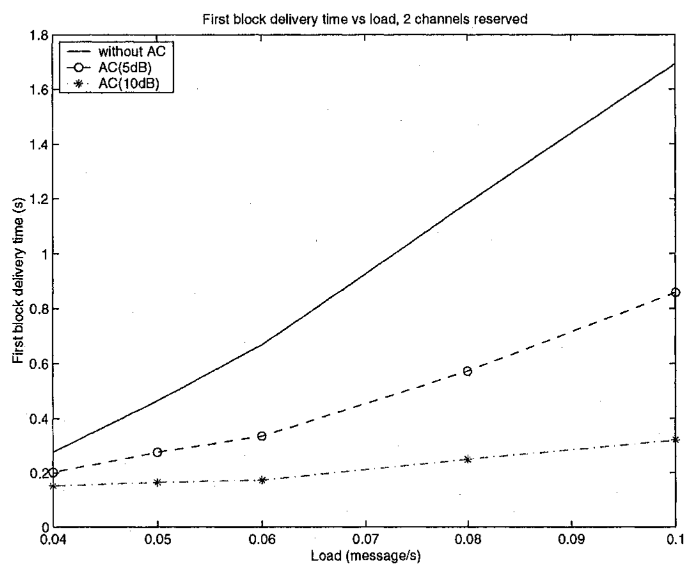


Figure D.22: First block delivery time , 0.06 call/s, call length 60 s

Appendix E

Results for QoS policy based on message size

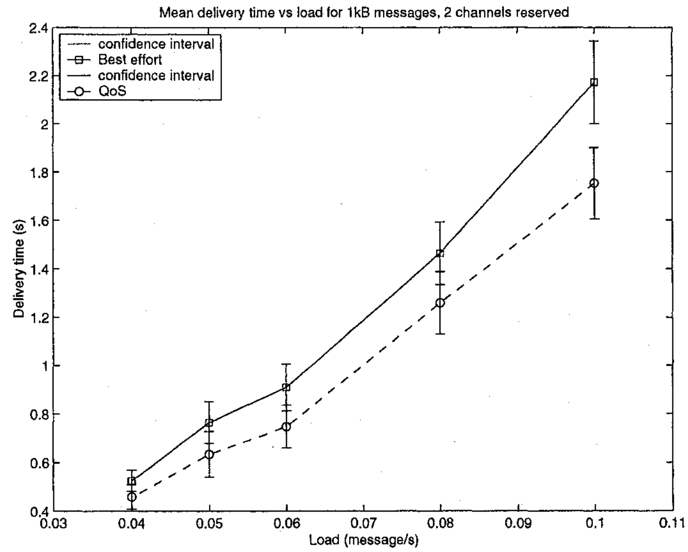


Figure E.1: Mean delivery time, 1KB messages, 0.06 call/s, call length 60 s

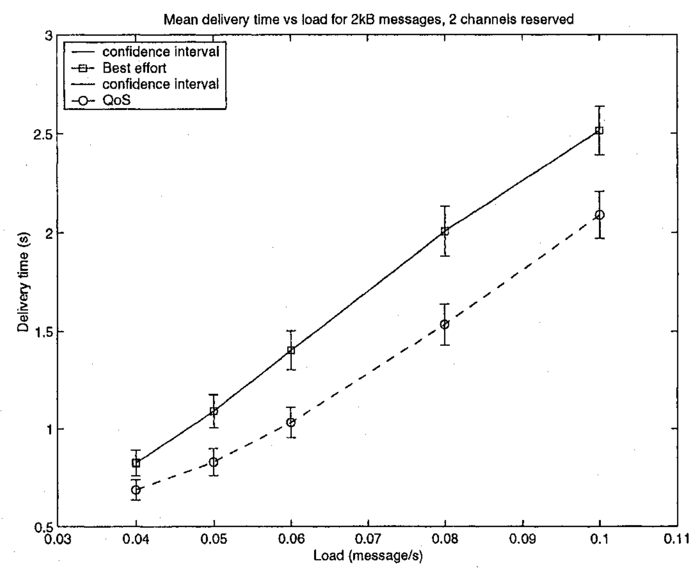


Figure E.2: Mean delivery time, 2KB messages, 0.06 call/s, call length 60 s

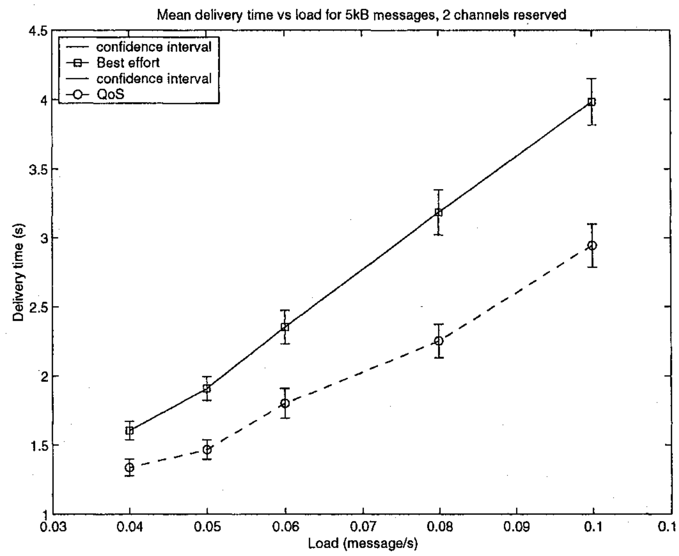


Figure E.3: Mean delivery time, 5KB messages, 0.06 call/s, call length 60 s

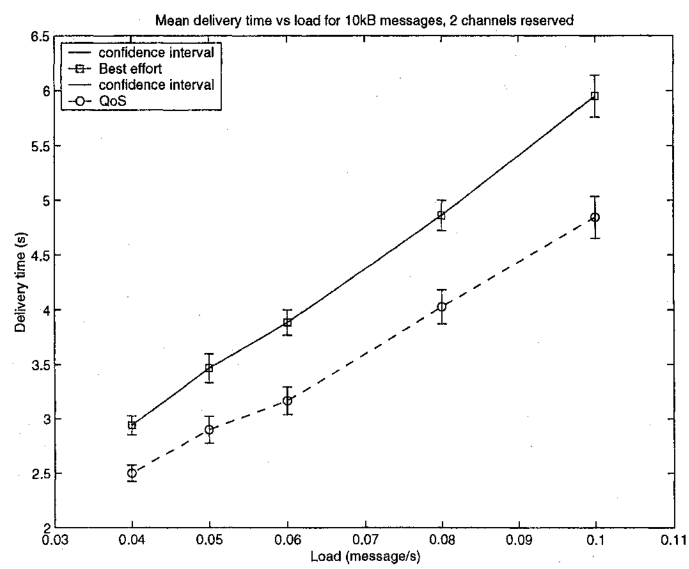


Figure E.4: Mean delivery time, 10KB messages, 0.06 call/s, call length 60 s

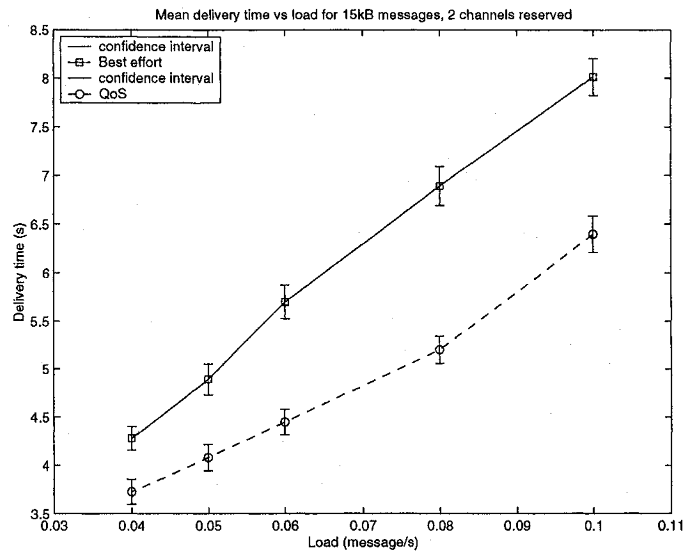


Figure E.5: Mean delivery time, 15KB messages, 0.06 call/s, call length 60 s

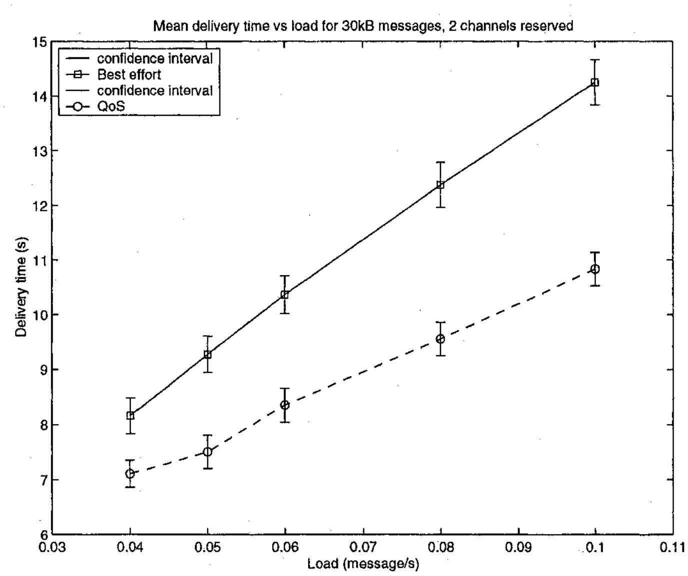


Figure E.6: Mean delivery time, 30KB messages, 0.06 call/s, call length 60 s

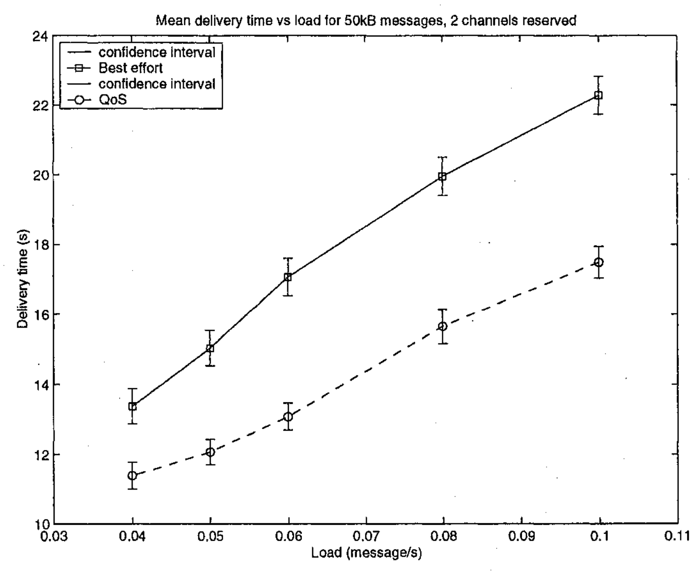


Figure E.7: Mean delivery time, 50KB messages, 0.06 call/s, call length 60 s

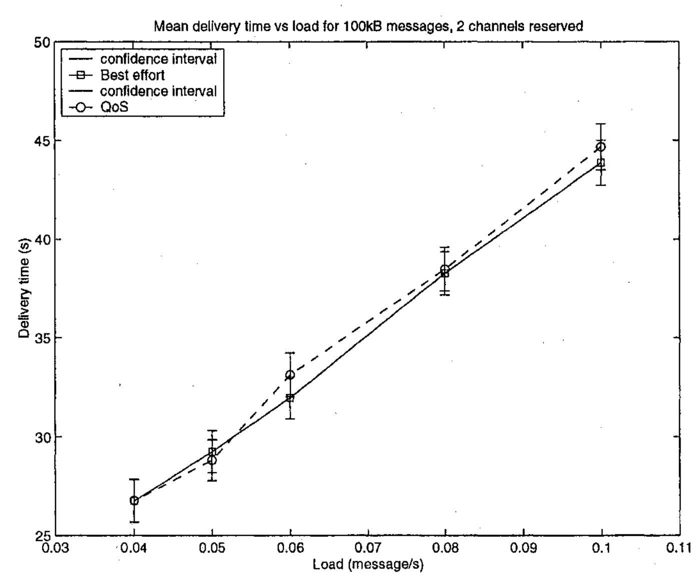


Figure E.8: Mean delivery time, 100KB messages, 0.06 call/s, call length 60 s

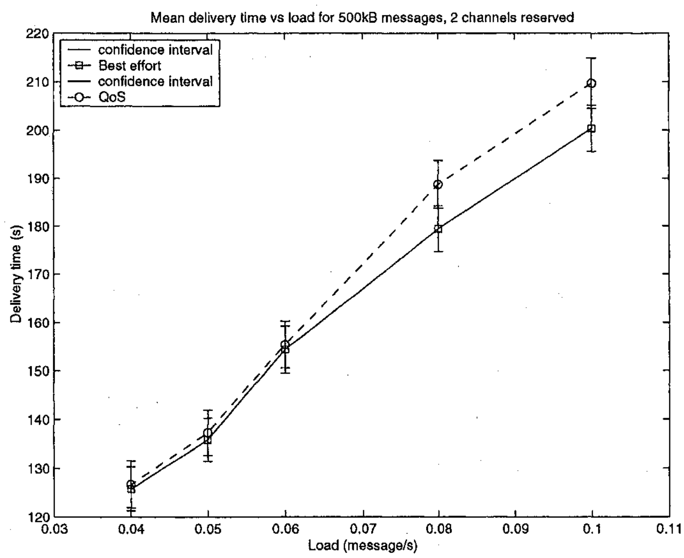


Figure E.9: Mean delivery time, 500KB messages, 0.06 call/s, call length 60 s

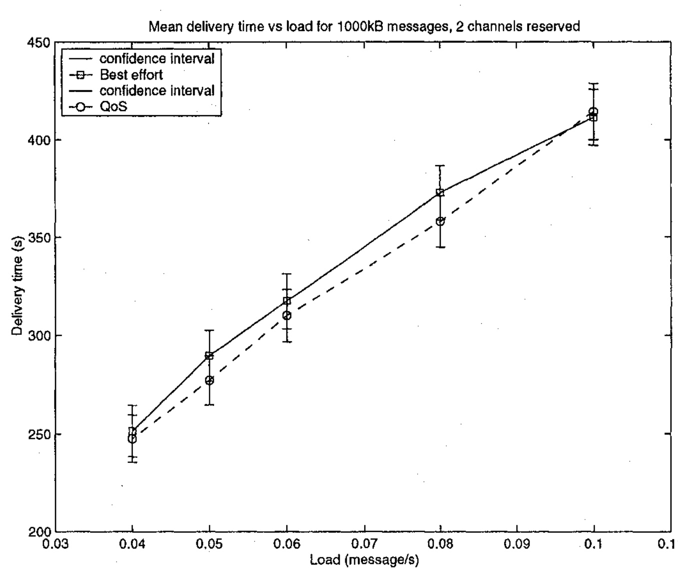


Figure E.10: Mean delivery time, 1000KB messages, 0.06 call/s, call length 60 s

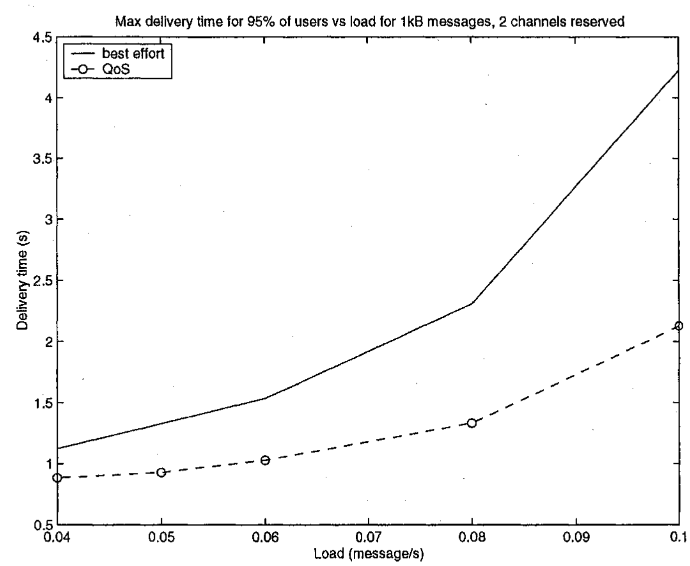


Figure E.11: Maximum delivery time for 95% of users, 1KB messages, 0.06 call/s, call length 60 s

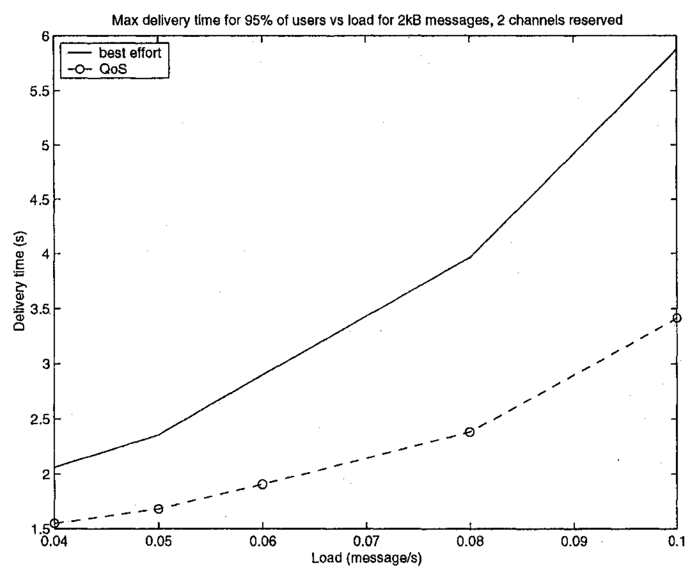


Figure E.12: Maximum delivery time for 95% of users, 2KB messages, 0.06 call/s, call length 60 s

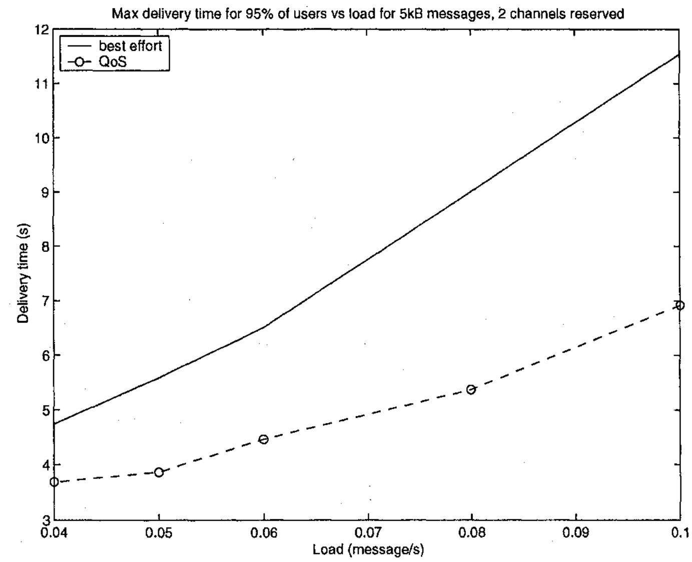


Figure E.13: Maximum delivery time for 95% of users, 5KB messages, 0.06 call/s, call length 60 s

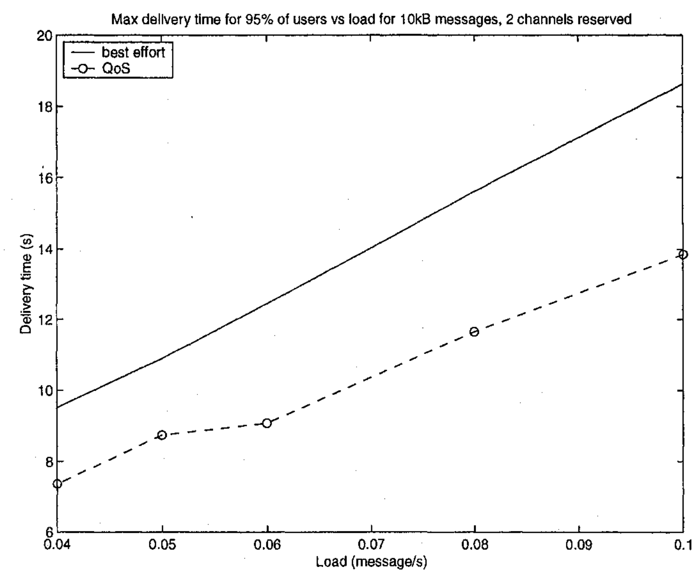


Figure E.14: Maximum delivery time for 95% of users, 10KB messages, 0.06 call/s, call length 60 s

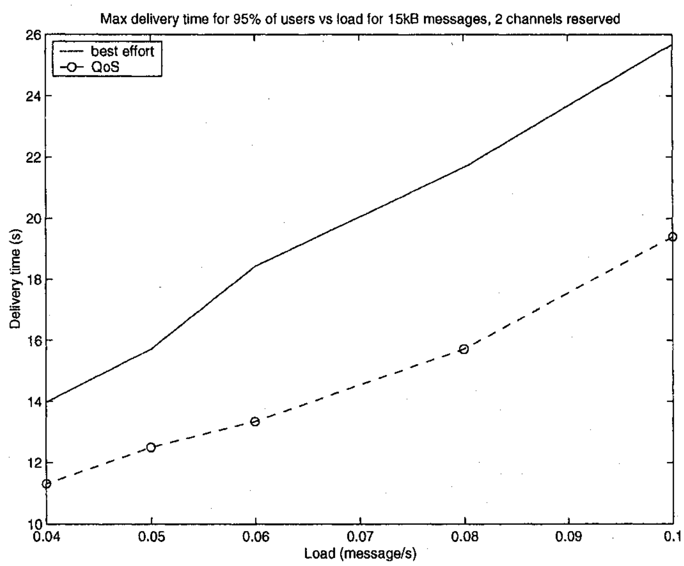


Figure E.15: Maximum delivery time for 95% of users, 15KB messages, 0.06 call/s, call length 60 s

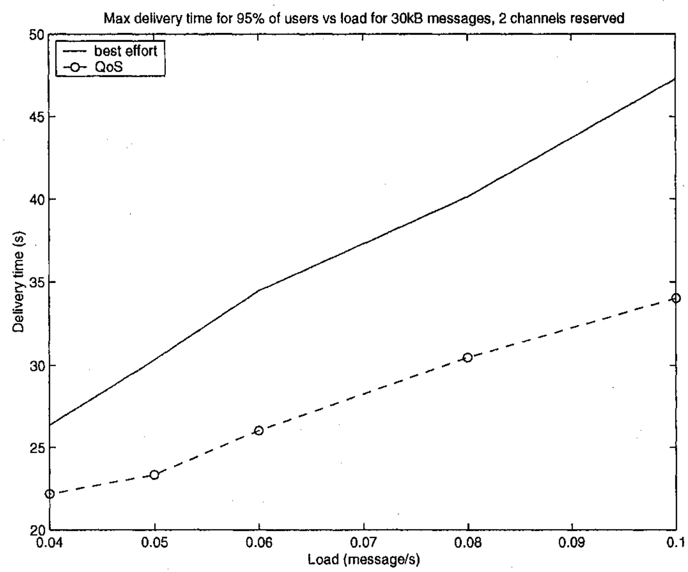


Figure E.16: Maximum delivery time for 95% of users, 30KB messages, 0.06 call/s, call length 60 s

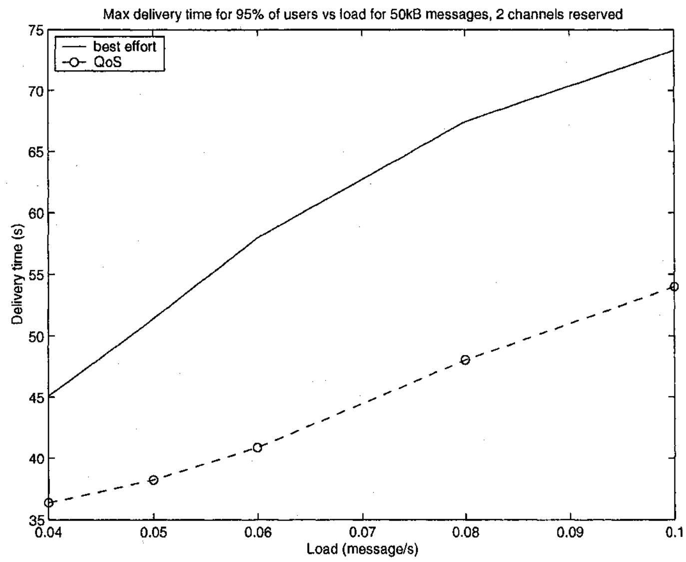


Figure E.17: Maximum delivery time for 95% of users, 50KB messages, 0.06 call/s, call length 60 s

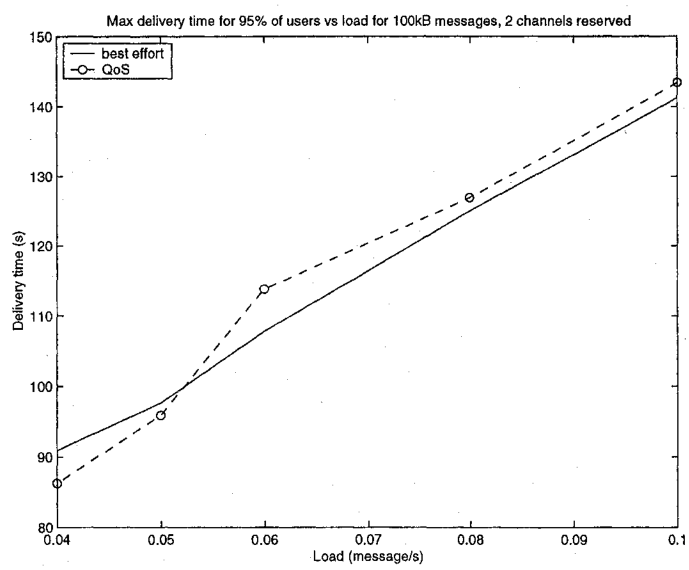


Figure E.18: Maximum delivery time for 95% of users, 100KB messages, 0.06 call/s, call length 60 s

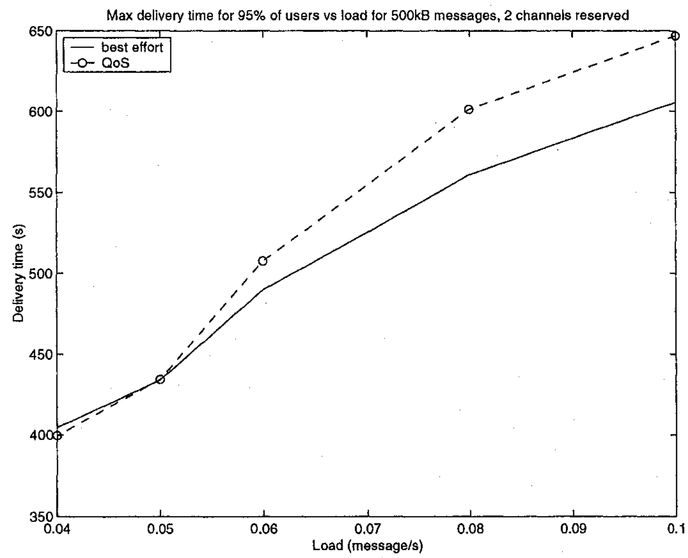


Figure E.19: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s

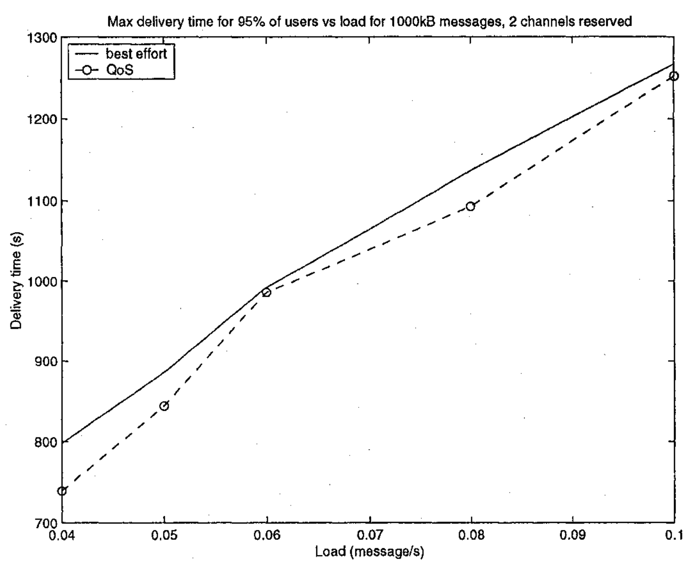


Figure E.20: Maximum delivery time for 95% of users, 1000KB messages, 0.06 call/s, call length 60 s

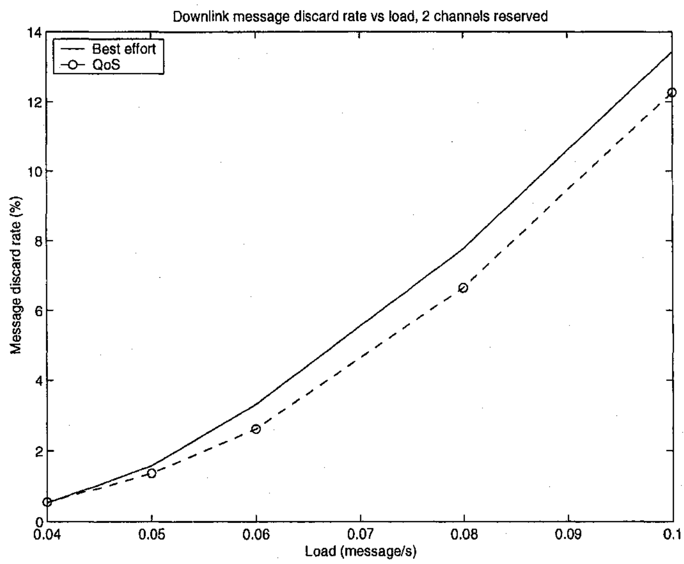


Figure E.21: Message discard rate, 0.06 call/s, call length 60 s

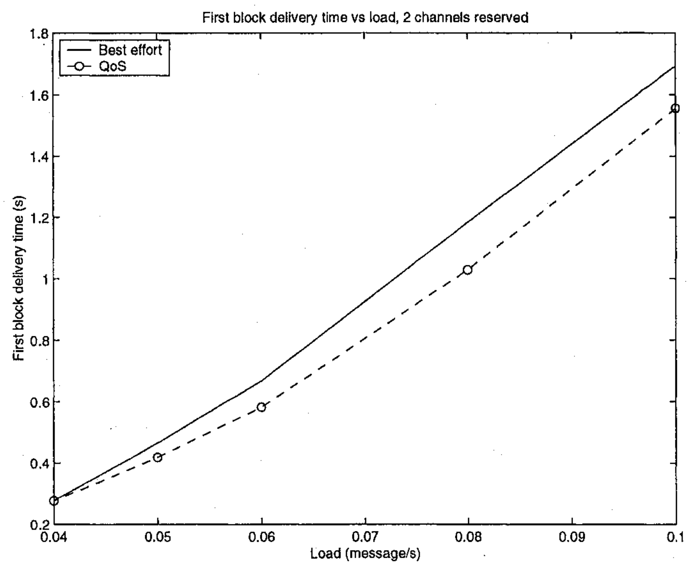


Figure E.22: First block delivery time , 0.06 call/s, call length 60 s

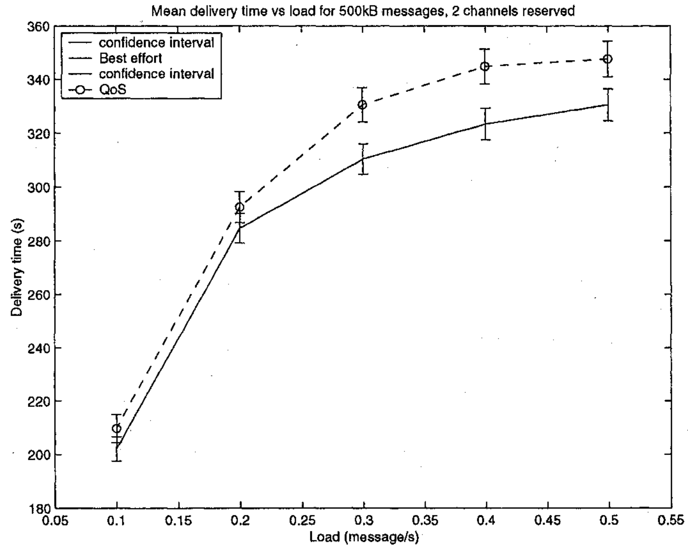


Figure E.23: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s

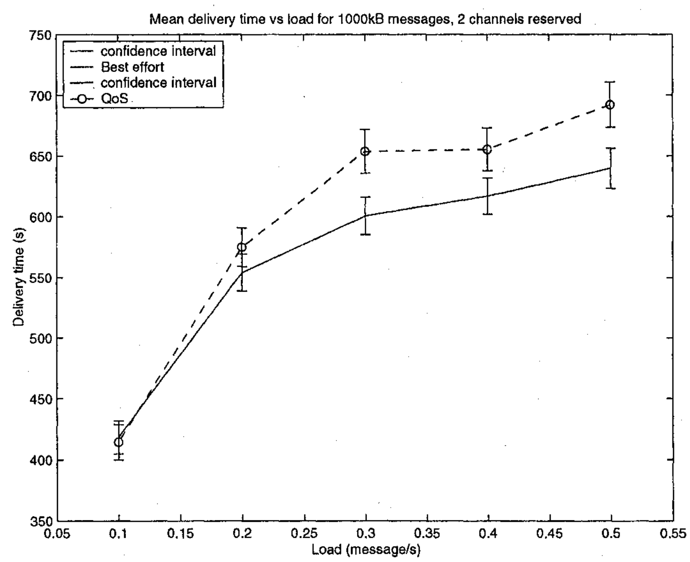


Figure E.24: Maximum delivery time for 95% of users, 1000KB messages, 0.06 call/s, call length 60 s

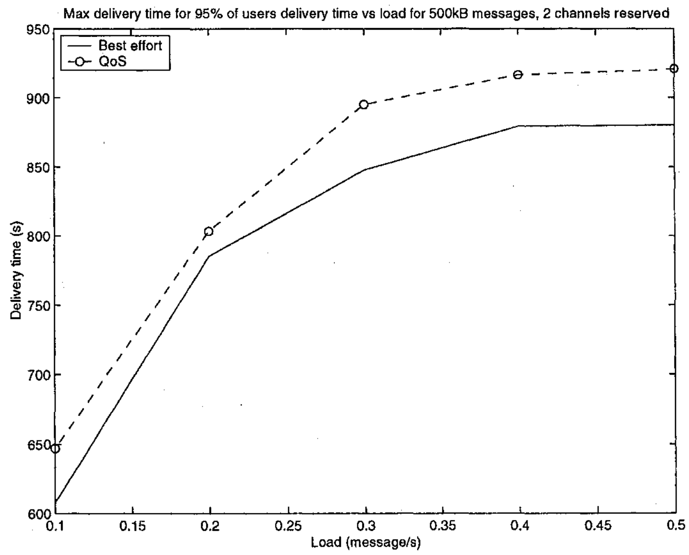


Figure E.25: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s

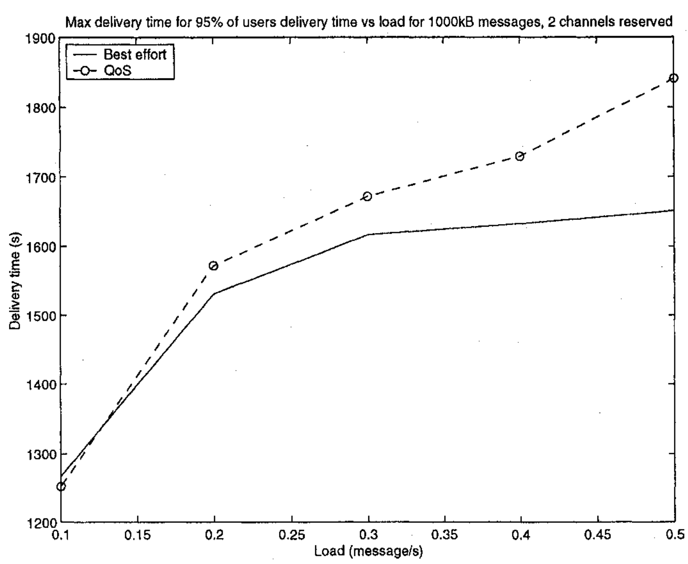


Figure E.26: Maximum delivery time for 95% of users, 1000KB messages, 0.06 call/s, call length 60 s

Appendix F

Results for QoS policy based on Differentiated Services

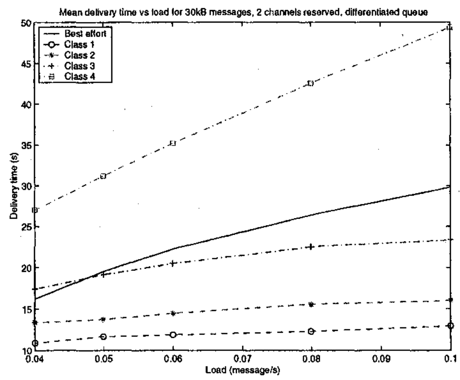


Figure F.1: Mean delivery time, 30KB messages, 0.06 call/s, call length 60 s, differentiated queue

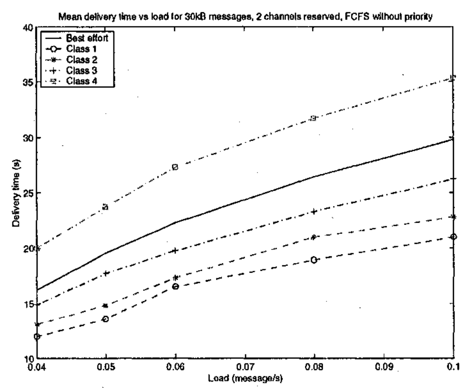


Figure F.2: Mean delivery time, 30KB messages, 0.06 call/s, call length 60 s, FCFS without priority

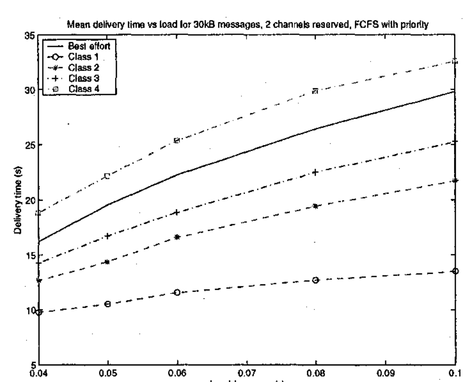


Figure F.3: Mean delivery time, 30KB messages, 0.06 call/s, call length 60s, FCFS with priority

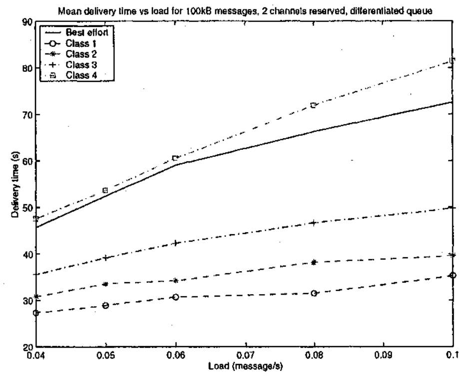


Figure F.4: Mean delivery time, 100KB messages, 0.06 call/s, call length 60 s, differentiated queue

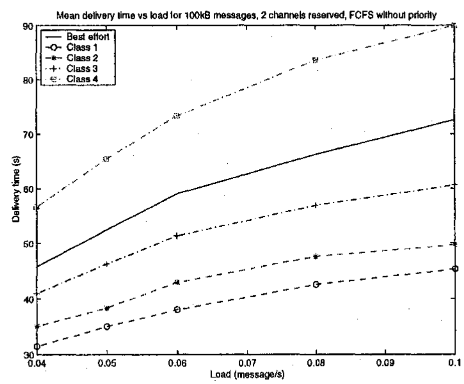


Figure F.5: Mean delivery time, 100KB messages, 0.06 call/s, call length 60 s, FCFS without priority

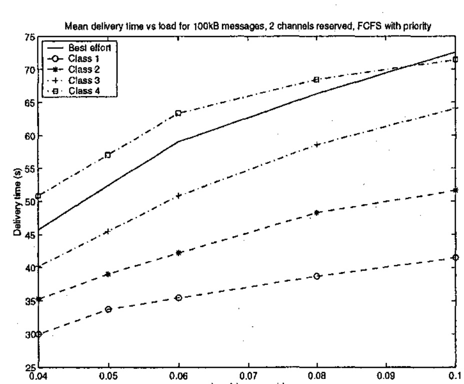


Figure F.6: Mean delivery time, 100KB messages, 0.06 call/s, call length 60 s, FCFS with priority

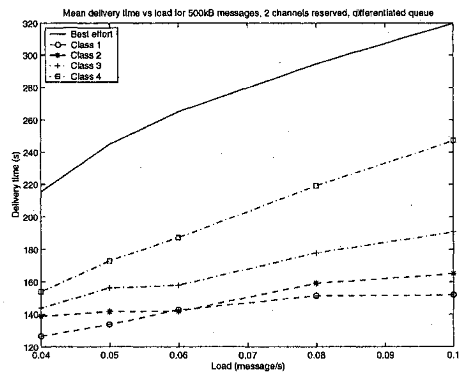


Figure F.7: Mean delivery time, 500KB messages, 0.06 call/s, call length 60 s, differentiated queue

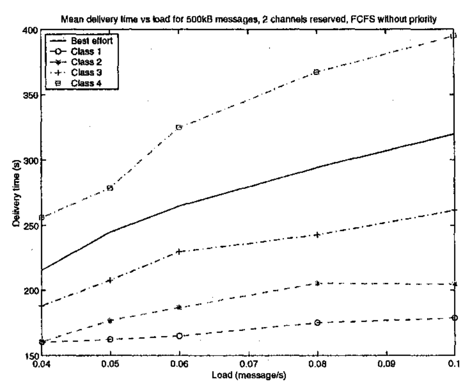


Figure F.8: Mean delivery time, 500KB messages, 0.06 call/s, call length 60 s, FCFS without priority

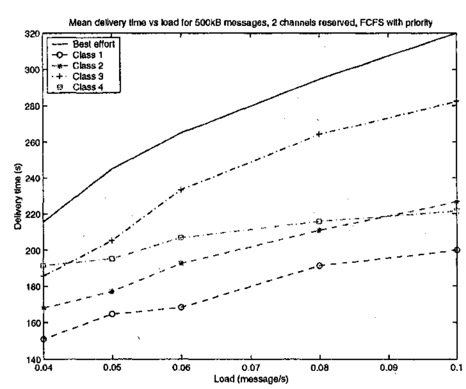


Figure F.9: Mean delivery time, 500KB messages, 0.06 call/s, call length 60 s, FCFS with priority

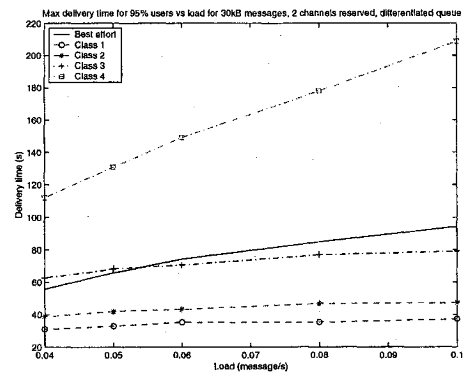


Figure F.10: Maximum delivery time for 95% of users, 30KB messages, 0.06 call/s, call length 60 s, differentiated queue

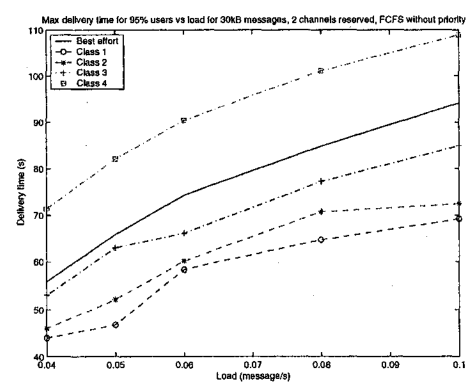


Figure F.11: Maximum delivery time for 95% of users, 30KB messages, 0.06 call/s, call length 60 s, FCFS without priority

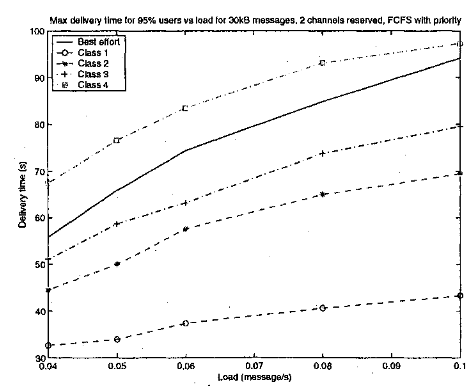


Figure F.12: Maximum delivery time for 95% of users, 30KB messages, 0.06 call/s, call length 60 s, FCFS with priority

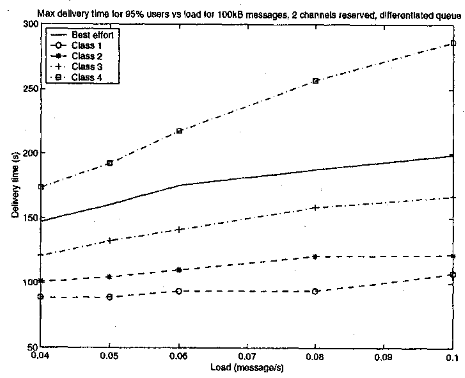


Figure F.13: Maximum delivery time for 95% of users, 100KB messages, 0.06 call/s, call length 60 s, differentiated queue

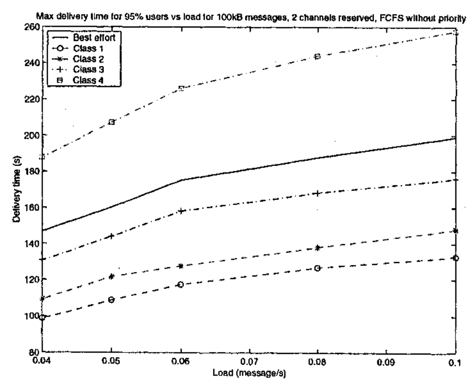


Figure F.14: Maximum delivery time for 95% of users, 100KB messages, 0.06 call/s, call length 60 s, FCFS without priority

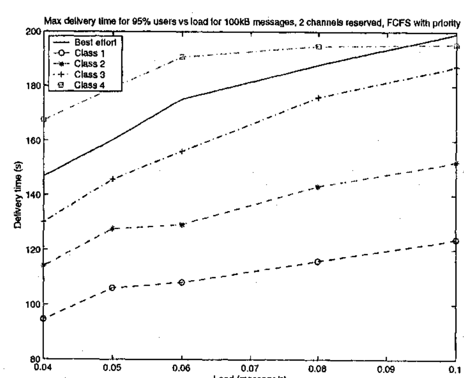


Figure F.15: Maximum delivery time for 95% of users, 100KB messages, 0.06 call/s, call length 60 s, FCFS with priority

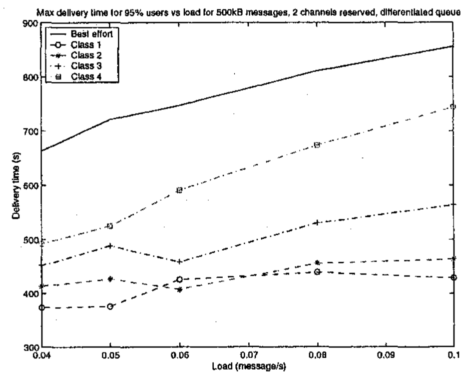


Figure F.16: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s, differentiated queue

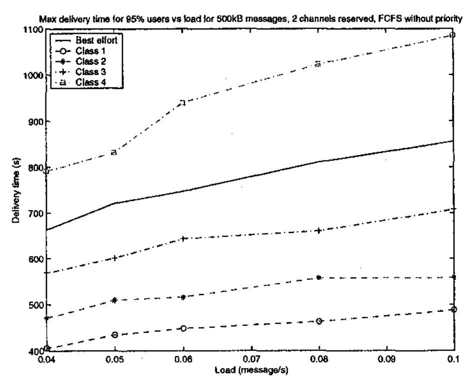


Figure F.17: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s, FCFS without priority

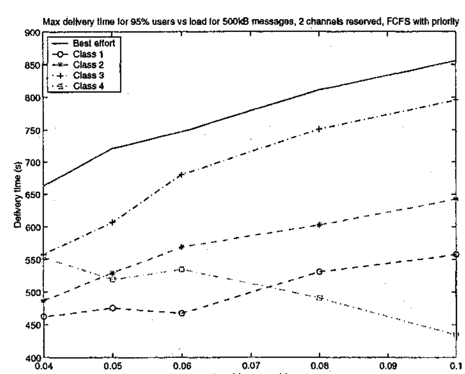


Figure F.18: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s, FCFS with priority

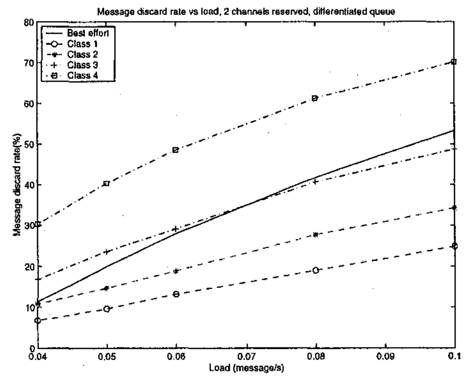


Figure F.19: Message discard rate, 0.06 call/s, call length 60 s, differentiated queue

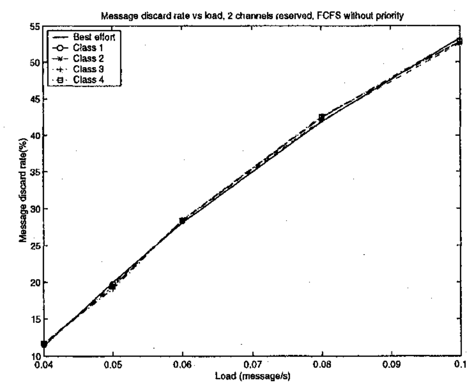


Figure F.20: Message discard rate, 0.06 call/s, call length 60 s, TBF queue without priority

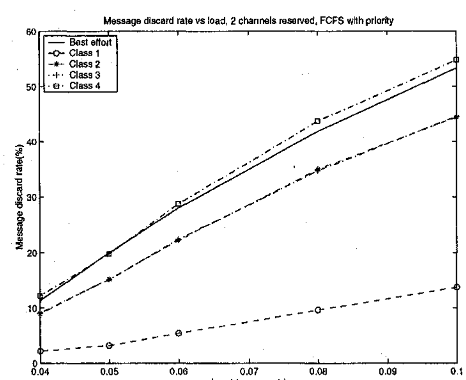


Figure F.21: Message discard rate, 0.06 call/s, call length 60 s, TBF queue with priority

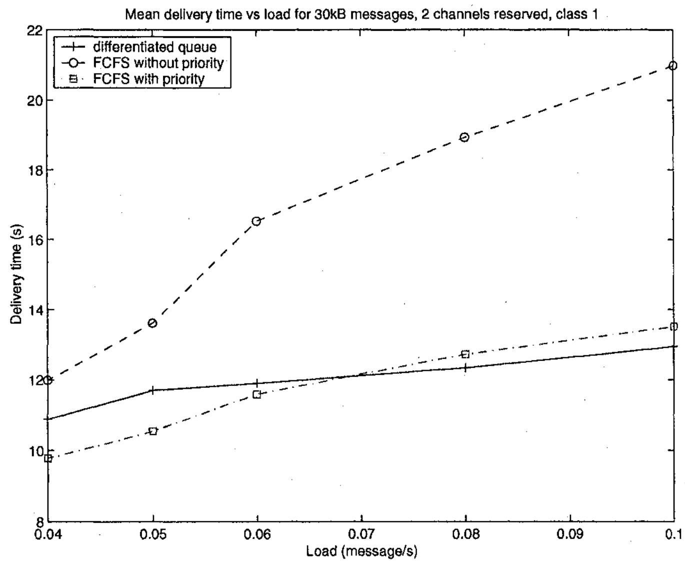


Figure F.22: Mean delivery time, 30KB messages, 0.06 call/s, call length 60 s, class 1

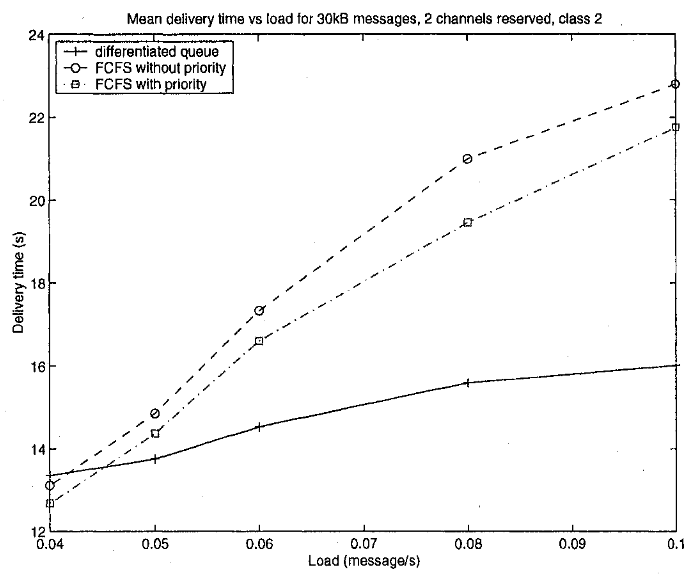


Figure F.23: Mean delivery time, 30KB messages, 0.06 call/s, call length 60 s, class 2

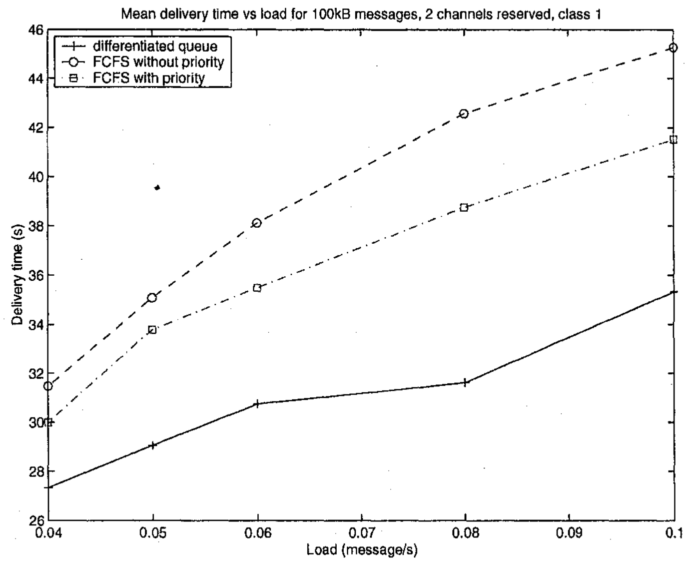


Figure F.24: Mean delivery time, 100KB messages, 0.06 call/s, call length 60 s, class 1

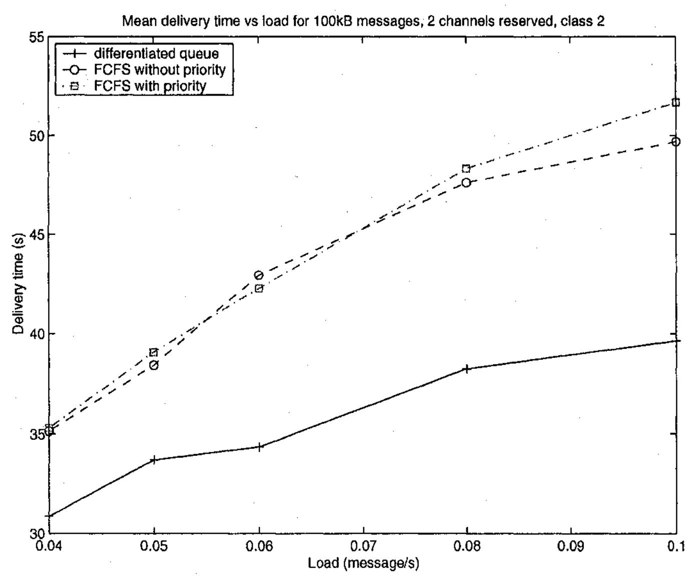


Figure F.25: Mean delivery time, 100KB messages, 0.06 call/s, call length 60 s, class 2

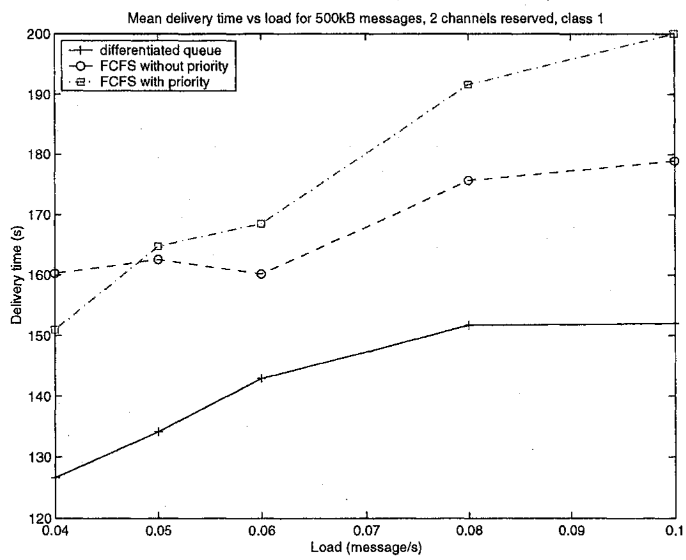


Figure F.26: Mean delivery time, 500KB messages, 0.06 call/s, call length 60 s, class 1

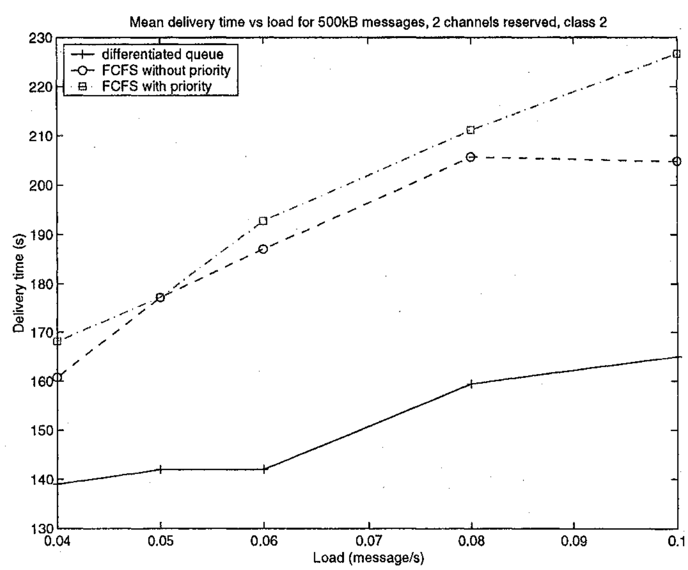


Figure F.27: Mean delivery time, 500KB messages, 0.06 call/s, call length 60 s, class 2

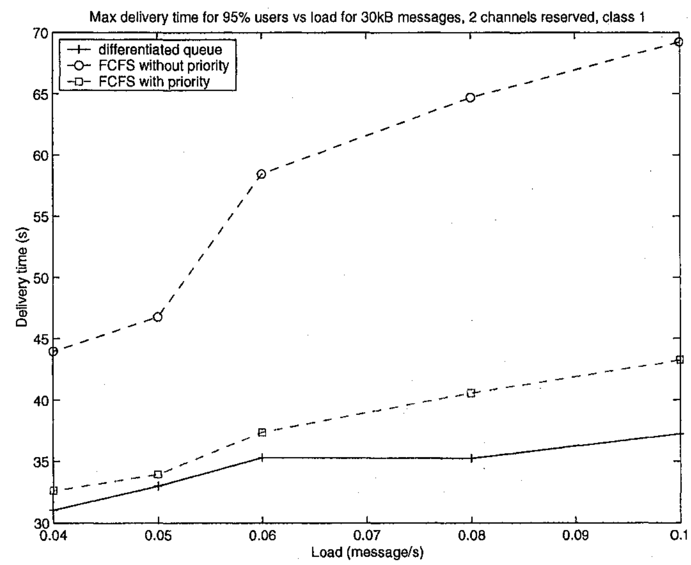


Figure F.28: Maximum delivery time for 95% of users, 30KB messages, 0.06 call/s, call length 60 s, class 1

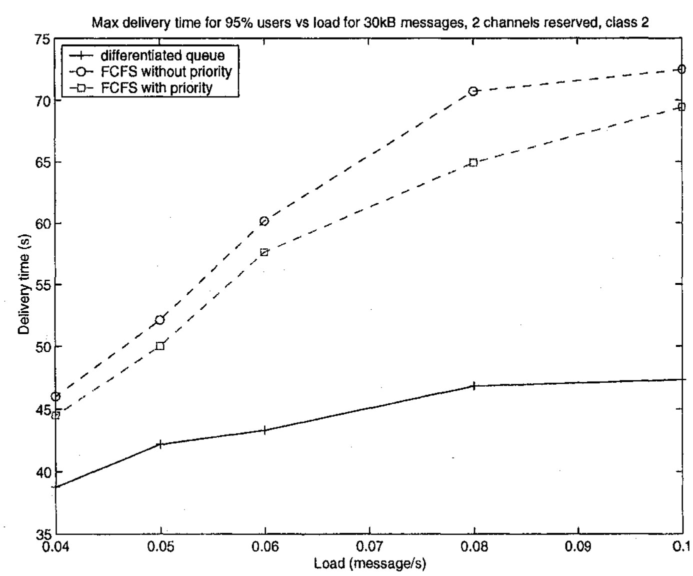


Figure F.29: Maximum delivery time for 95% of users, 30KB messages, 0.06 call/s, call length 60 s, class 2

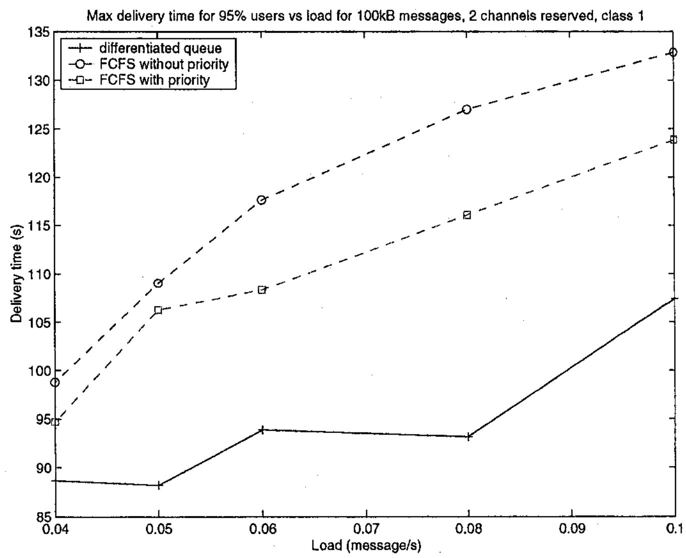


Figure F.30: Maximum delivery time for 95% of users, 100KB messages, 0.06 call/s, call length 60 s, class 1

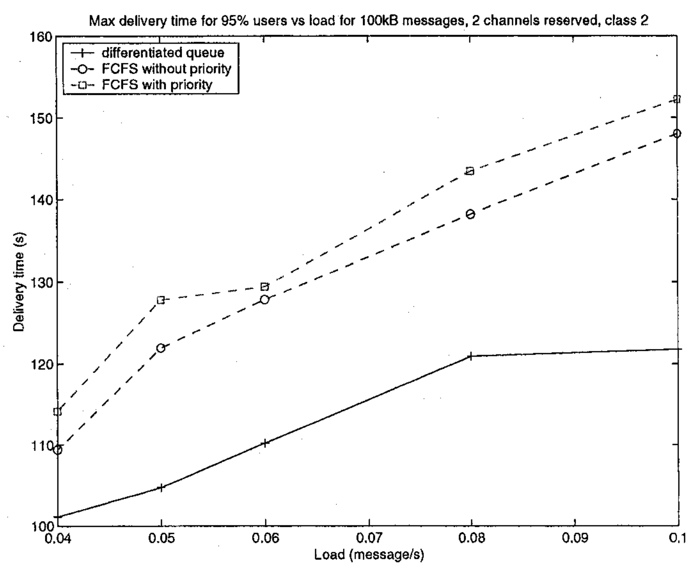


Figure F.31: Maximum delivery time for 95% of users, 100KB messages, 0.06 call/s, call length 60 s, class 2

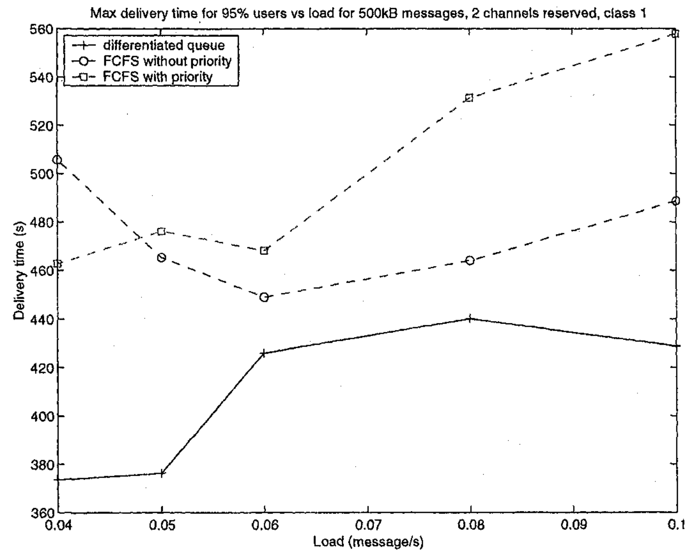


Figure F.32: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s, class 1

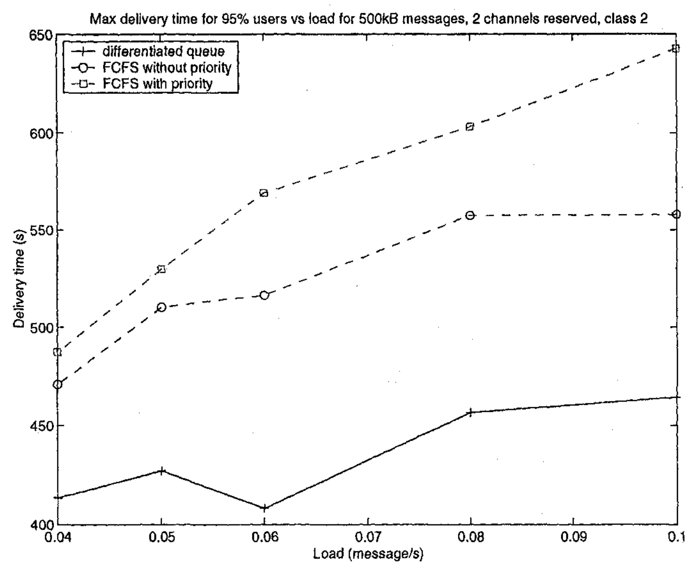


Figure F.33: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s, class 2

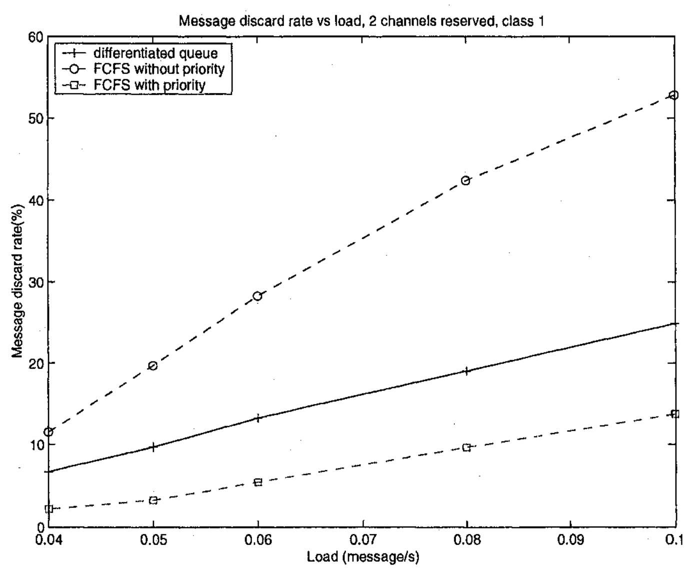


Figure F.34: Message discard rate, 0.06 call/s, call length 60 s, class 1

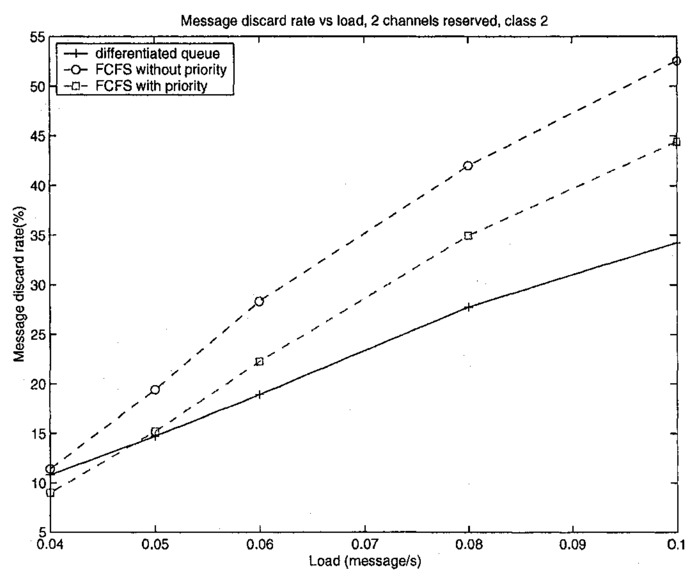


Figure F.35: Message discard rate, 0.06 call/s, call length 60 s, class 2

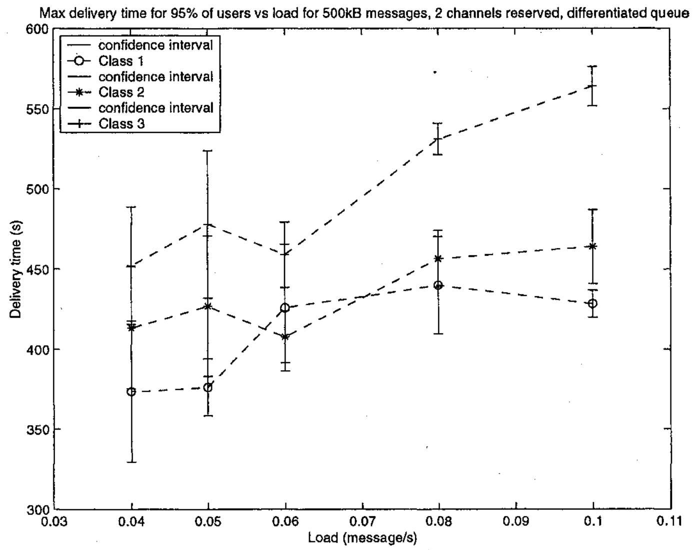


Figure F.36: Maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s, differentiated queue

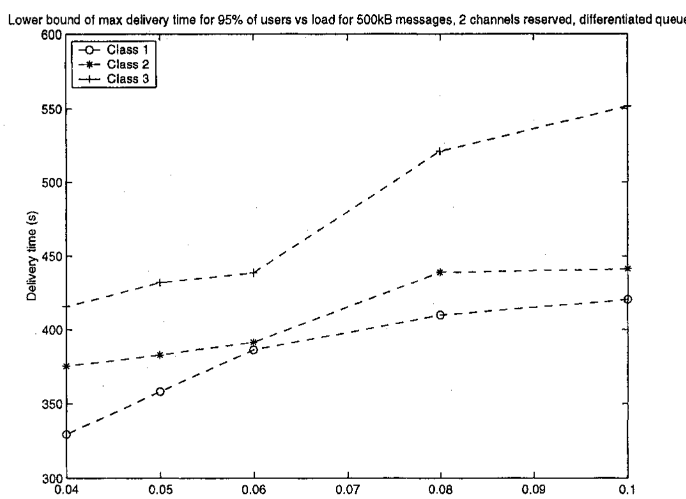


Figure F.37: Lower bound of maximum delivery time for 95% of users, 500KB messages, 0.06 call/s, call length 60 s, differentiated queue